

NORTH ATLANTIC TREATY ORGANIZATION



RESEARCH AND TECHNOLOGY ORGANIZATION

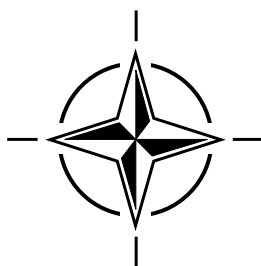
BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO MEETING PROCEEDINGS 32

The Human Factor in System Reliability – Is Human Performance Predictable?

(les Facteurs humains et la fiabilité des systèmes –
Les performances humaines, sont-elles prévisibles?)

*Papers presented at the Human Factors and Medicine Panel (HFM) Workshop held in Siena, Italy
from 1-2 December 1999.*



Published January 2001

Distribution and Availability on Back Cover

Form SF298 Citation Data

Report Date <i>("DD MON YYYY")</i> 01012001	Report Type N/A	Dates Covered (from... to) <i>("DD MON YYYY")</i>
Title and Subtitle The Human Factor in System Reliability Is Human Performance Predictable?		Contract or Grant Number
Authors		Program Element Number
		Project Number
		Task Number
Performing Organization Name(s) and Address(es) Research and Technology Organization North Atlantic Treaty Organization BP 25, 7 rue Ancelle F92201 Neuilly-sur-Seine Cedex, France		Work Unit Number
		Performing Organization Number(s)
		Monitoring Agency Name(s) and Address(es)
Sponsoring/Monitoring Agency Name(s) and Address(es)		Monitoring Agency Acronym
		Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited		
Supplementary Notes		
Abstract Human error is seen as an unacceptably high contributing factor in most military accidents and much research has been carried out over the past 50 years, to attempt to predict the probability of the occurrence of human error. Significant advances have been made within the safety critical domain areas within the nuclear and chemical industries. The aim of the workshop was to review the research carried out across multiple domain areas in order to provide a clear focus for Working Group 30 (Human Reliability in Safety Critical Systems). It was evident from the workshop that key cognitive processes and organisational contexts play an important part in shaping the overall human performance and hence the likelihood of human error. Therefore it was clear that there are new approaches to Human Reliability Assessment that take account of the unique human adaptability attributes that are not present in any other part of the overall system in which the human is an integral part. Working Group 30 will develop these approaches to provide clear guidance to the NATO community in designing and analysing human roles to quantify and qualify the likelihood of error. This will enhance future design processes to produce higher fault tolerant designs, to include mitigating strategies and aim towards a significant reduction in the number of human errors.		

Subject Terms

Human factors engineering; Cognition; Adaptation; Performance; Systems analysis; Design; Reliability; Aviation safety; Procedures; Safety; Predictions; Organizations; Errors; Accident investigations; Models; Performance evaluation; Humans; Human behavior

Document Classification

unclassified

Classification of SF298

unclassified

Classification of Abstract

unclassified

Limitation of Abstract

unlimited

Number of Pages

111

REPORT DOCUMENTATION PAGE																					
1. Recipient's Reference	2. Originator's References RTO-MP-032 AC/323(HFM)TP/12	3. Further Reference ISBN 92-837-1053-3	4. Security Classification of Document UNCLASSIFIED/ UNLIMITED																		
5. Originator	Research and Technology Organization North Atlantic Treaty Organization BP 25, 7 rue Ancelle, F-92201 Neuilly-sur-Seine Cedex, France																				
6. Title	The Human Factor in System Reliability – Is Human Performance Predictable?																				
7. Presented at/sponsored by	the Human Factors and Medicine Panel (HFM) Workshop held in Siena, Italy from 1-2 December 1999.																				
8. Author(s)/Editor(s) Multiple			9. Date January 2001																		
10. Author's/Editor's Address Multiple			11. Pages 110																		
12. Distribution Statement	There are no restrictions on the distribution of this document. Information about the availability of this and other RTO unclassified publications is given on the back cover.																				
13. Keywords/Descriptors																					
<table border="0"> <tbody> <tr> <td>Human factors engineering</td> <td>Cognition</td> <td>Adaptation</td> </tr> <tr> <td>Performance</td> <td>Systems analysis</td> <td>Design</td> </tr> <tr> <td>Reliability</td> <td>Aviation safety</td> <td>Procedures</td> </tr> <tr> <td>Safety</td> <td>Predictions</td> <td>Organizations</td> </tr> <tr> <td>Errors</td> <td>Accident investigations</td> <td>Models</td> </tr> <tr> <td>Performance evaluation</td> <td>Humans</td> <td>Human behavior</td> </tr> </tbody> </table>				Human factors engineering	Cognition	Adaptation	Performance	Systems analysis	Design	Reliability	Aviation safety	Procedures	Safety	Predictions	Organizations	Errors	Accident investigations	Models	Performance evaluation	Humans	Human behavior
Human factors engineering	Cognition	Adaptation																			
Performance	Systems analysis	Design																			
Reliability	Aviation safety	Procedures																			
Safety	Predictions	Organizations																			
Errors	Accident investigations	Models																			
Performance evaluation	Humans	Human behavior																			
14. Abstract																					
<p>Human error is seen as an unacceptably high contributing factor in most military accidents and much research has been carried out over the past 50 years, to attempt to predict the probability of the occurrence of human error. Significant advances have been made within the safety critical domain areas within the nuclear and chemical industries. The aim of the workshop was to review the research carried out across multiple domain areas in order to provide a clear focus for Working Group 30 (Human Reliability in Safety Critical Systems). It was evident from the workshop that key cognitive processes and organisational contexts play an important part in shaping the overall human performance and hence the likelihood of human error. Therefore it was clear that there are new approaches to Human Reliability Assessment that take account of the unique human adaptability attributes that are not present in any other part of the overall system in which the human is an integral part. Working Group 30 will develop these approaches to provide clear guidance to the NATO community in designing and analysing human roles to quantify and qualify the likelihood of error. This will enhance future design processes to produce higher fault tolerant designs, to include mitigating strategies and aim towards a significant reduction in the number of human errors.</p>																					

This page has been deliberately left blank



Page intentionnellement blanche

NORTH ATLANTIC TREATY ORGANIZATION



RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO MEETING PROCEEDINGS 32

The Human Factor in System Reliability – Is Human Performance Predictable?

(les Facteurs humains et la fiabilité des systèmes – Les performances humaines, sont-elles prévisibles?)

Papers presented at the Human Factors and Medicine Panel (HFM) Workshop held in Siena, Italy from 1-2 December 1999.



The Research and Technology Organization (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote cooperative research and information exchange. The objective is to support the development and effective use of national defence research and technology and to meet the military needs of the Alliance, to maintain a technological lead, and to provide advice to NATO and national decision makers. The RTO performs its mission with the support of an extensive network of national experts. It also ensures effective coordination with other NATO bodies involved in R&T activities.

RTO reports both to the Military Committee of NATO and to the Conference of National Armament Directors. It comprises a Research and Technology Board (RTB) as the highest level of national representation and the Research and Technology Agency (RTA), a dedicated staff with its headquarters in Neuilly, near Paris, France. In order to facilitate contacts with the military users and other NATO activities, a small part of the RTA staff is located in NATO Headquarters in Brussels. The Brussels staff also coordinates RTO's cooperation with nations in Middle and Eastern Europe, to which RTO attaches particular importance especially as working together in the field of research is one of the more promising areas of initial cooperation.

The total spectrum of R&T activities is covered by the following 7 bodies:

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS Studies, Analysis and Simulation Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These bodies are made up of national representatives as well as generally recognised 'world class' scientists. They also provide a communication link to military users and other NATO bodies. RTO's scientific and technological work is carried out by Technical Teams, created for specific activities and with a specific duration. Such Technical Teams can organise workshops, symposia, field trials, lecture series and training courses. An important function of these Technical Teams is to ensure the continuity of the expert networks.

RTO builds upon earlier cooperation in defence research and technology as set-up under the Advisory Group for Aerospace Research and Development (AGARD) and the Defence Research Group (DRG). AGARD and the DRG share common roots in that they were both established at the initiative of Dr Theodore von Kármán, a leading aerospace scientist, who early on recognised the importance of scientific support for the Allied Armed Forces. RTO is capitalising on these common roots in order to provide the Alliance and the NATO nations with a strong scientific and technological basis that will guarantee a solid base for the future.

The content of this publication has been reproduced directly from material supplied by RTO or the authors.

Published January 2001

Copyright © RTO/NATO 2001
All Rights Reserved

ISBN 92-837-1053-3



*Printed by St. Joseph Ottawa/Hull
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada J8X 1C6*

The Human Factor in System Reliability – Is Human Performance Predictable?

(RTO MP-032)

Executive Summary

This workshop was convened by the Human Factors and Medicine (HFM) Panel of the Research and Technology Organisation (RTO) mainly as a precursor to a new Task Group WG30 which aims to investigate and develop the role of Human Reliability Assessment within the overall design process. The workshop attracted an excellent collection of experts and practitioners across both civil and military domains and was successful in highlighting the needs of the Human Reliability community and hence providing a clear focus for the newly formed Task Group.

The workshop received eleven papers, including two keynote addresses, which tackled a gamut of issues including:

- Current Safety Assessment methods
- Limitations of existing Human Performance Models
- Cognitive Reliability Analysis techniques
- Barrier functions and their impact on human reliability
- Quantitative vs Qualitative analytical approaches
- Characteristics of high reliability organisations
- Variability of Corporate Safety Cultures
- Contextual Causal Modelling Techniques
- Cost Effectiveness Analysis techniques in Human Reliability Modelling
- Causal Database Developments
- Application of Hierarchical Taxonomy approaches to Human Error Prediction

The workshop identified and debated recent trends in Human Reliability Assessment, in particular the pressure to treat human error analysis in the same manner as hardware component analysis. The need for new techniques in data collection, data analysis and human error quantification was examined that take account of unique human attributes.

New themes that emerged included a need to focus more upon cognitive processes and the organisational context in which system behaviour takes place. The traditional focus on human error should be broadened to consider human adaptability as a safety feature and the classical concept of a human task expanded to embrace a wider work scenario. The concept of high reliability cultures and organisation was also addressed with emphasis being placed on the development and adoption of proactive safe working practices.

In conclusion, the workshop was highly successful in sharing and debating state of the art knowledge and assessment approaches that will collectively enhance the science of human reliability within the overall design process. The valuable theoretical and practical insights contained in the presentations coupled with the lively debate on the issues raised, ensured that the workshop not only fulfilled its objectives from an educational standpoint, but also provided an enjoyable yet formative experience for the delegates.

les Facteurs humains et la fiabilité des systèmes – Les performances humaines, sont-elles prévisibles?

(RTO MP-032)

Synthèse

Cet atelier a été organisé par la commission sur les facteurs humains et la médecine (HFM) de l'Organisation pour la recherche et la technologie de l'OTAN (RTO), comme précurseur au nouveau groupe de travail WG30, dont l'objectif est d'examiner et de développer le rôle de l'«Evaluation de la Fiabilité Humaine» dans le processus de conception. De nombreux spécialistes et praticiens civils et militaires ont été attirés par le sujet de cet atelier, qui a permis de mettre en évidence les besoins des chercheurs travaillant dans le domaine de la fiabilité humaine et par conséquent, d'établir une base de travail claire pour le nouveau groupe de travail.

Onze communications, dont deux discours d'ouverture, ont été présentées lors de l'atelier, qui a permis d'examiner un grand éventail de questions dont les suivantes :

- Méthodes actuelles d'évaluation de la sécurité
- Limitations des modèles actuels de performances humaines
- Techniques d'analyse de la fiabilité cognitive
- Fonctions barrière et leur impact sur la fiabilité humaine
- Approches analytiques quantitatives contre approches analytiques qualitatives
- Caractéristiques des organisations hautement fiables
- Variabilité des cultures sur la sécurité dans l'entreprise
- Techniques contextuelles de modélisation causale
- Techniques d'analyse coût-efficacité dans la modélisation de la fiabilité humaine
- Développements dans le domaine des bases de données causales
- Application d'approches basées sur la taxonomie hiérarchique à la prévision de l'erreur humaine

L'atelier a permis d'identifier et de discuter des tendances récentes dans le domaine de l'évaluation de la fiabilité humaine et en particulier la tendance de plus en plus marquée qui veut que l'on traite l'analyse de l'erreur humaine de la même manière que l'analyse des composants matériels. Le besoin de nouvelles techniques de collecte de données, de quantification de l'erreur humaine et d'analyse de données, qui tiendraient compte des attributs spécifiques à l'être humain, a aussi été abordé.

Le besoin de privilégier les processus cognitifs et le contexte organisationnel dans lequel les systèmes fonctionnent sont des exemples de nouveaux thèmes qui ont été mis en évidence pendant l'atelier. Le champ d'investigation classique de l'erreur humaine doit être élargi pour englober l'adaptabilité humaine en tant que facteur de sécurité; de plus, le concept traditionnel de la tâche humaine doit être étendu pour englober des scénarios de travail plus diversifiés. Le concept de cultures et d'organisations de haute fiabilité a également été examiné, l'accent étant mis sur le développement et l'adoption de pratiques de travail proactives sans danger.

En conclusion, l'atelier a largement permis de mettre en commun et de débattre les dernières approches de l'évaluation des connaissances, qui sont susceptibles de faire avancer la science de la fiabilité humaine dans le processus global de conception. Les précieux éclaircissements théoriques et pratiques contenus dans les présentations, associés aux vifs débats qui ont animé l'atelier ont permis non seulement aux organisateurs d'atteindre leurs objectifs pédagogiques, mais aussi à l'assistance de participer à une manifestation à la fois agréable et formatrice.

Contents

	Page
Executive Summary	iii
Synthèse	iv
Human Factors and Medicine Panel	vi
Technical Evaluation Report by T. Kontogiannis	Reference T
Keynote Address 1: Anticipating Failures: What Should Predictions Be About? by E. Hollnagel	KN1
SESSION I: CAN HUMAN PERFORMANCE BE ADDRESSED WITHIN THE CURRENT SAFETY ASSESSMENT PROCESS?	
Can Human Performance be Addressed Within the Current Safety Assessment Process? by M. Boasson	1
SESSION II: CAN IT BE PREDICTED? QUANTITATIVE AND QUALITATIVE ASPECTS TOOLS AND TECHNIQUES	
THEA – A Technique for Human Error Assessment Early in Design by S. Pocock, P. Wright and M. Harrison	2
Human Reliability in Civil Aircraft Inspection by C.G. Drury	3
Keynote Address 2: Impact of Organisational Factors on Effective Human Reliability Assessment by J. Reason	KN2
SESSION III: HOW ARE COGNITIVE FACTORS ADDRESSED IN SYSTEM RELIABILITY?	
Addressing Cognitive Factors in System Reliability by N. Moray	4
Effects of Practice and Memory Aiding on Decision Performance and Information Search in Command and Control by P.H.M.P. Roelofsma	5
SESSION IV: DATA COLLECTION: QUANTITATIVE AND QUALITATIVE ASPECTS	
The Risk of Human Error: Data Collection, Collation, and Quantification by J.W. Chappelow	6
Causal Models of Human Error in Accident Investigation: the Link Between Prediction and Prevention by D. Embrey	7
SESSION V: ORGANISATIONAL DIMENSIONS OF HUMAN RELIABILITY	
Safety Culture – Theory and Practice by P. Hudson	8
SHELFS: A Proactive Method for Managing Safety Issues by A. Rizzo and L. Save	9

Human Factors and Medicine Panel

Chairman:

Dr M.C. WALKER

Director, Centre for Human Sciences
DERA
F138 Building - Room 204
Farnborough, Hants GU14 0LX
United Kingdom

Co-Chairman:

Col. W.C.M. TIELEMANS, MD

RNLAF/SGO
P.O. Box 20703
Binckhorstlaan, 135
2500 ES The Hague
The Netherlands

WORKSHOP PROGRAMME COMMITTEE

Chairman

Dr D. EMBREY

Human Reliability Associates
1 School House
Higher Lane
Dalton, Wigan, Lancashire WN8 7RP, UK
dembrey@humanreliability.com

Members

Dipl.-Ing. F. FLEMISCH

University of Armed Forces Munich
Institut for System Dynamics and Flight Mechanics
Werner-Heisenbergweg 39
D-85577 Neubiberg, Germany
frank.flemisch@unibw-muenchen.de

DR W. KAEPLER

FGA-FKIE
Neuenahrer Strasse 20
D-53343 Wachtberg-Werthhoven, Germany
kaeppler@fgan.de

Dr R. ONKEN

University of Armed Forces Munich
Institut for System Dynamics and Flight Mechanics
Werner-Heisenbergweg 39
D-85577 Neubiberg, Germany
reiner.onken@unibw-muenchen.de

Dr W. JANSSEN

TNO Human Factors
Kampweg 5
NL-3769 DE Soesterberg, The Netherlands
janssen@tm.tno.nl

LtCol K. TUNGESVIK

HQS Defence Command Norway
Oslo Mil/Huseby
N-0016 Oslo, Norway
Tel: +47 23098797

Dr A. AKIN

GATA Have Uzay Hekimligi
Hava Hastanesi
26020 Eskisehir, Turkey
gang@marketweb.net.tr

Mr S. HARDING

Maritime and Coastguard Agency
MSOS (A) Bay 2/04
Spring Place
105 Commercial Road
Southampton, SO15 1EG, UK
steve_harding@mca.gov.uk

Dr P. WILKINSON

British Aerospace MA&A
Warton Aerodrome, W392A
Preston, Lancs PR4 1AX, UK
peter.r.wilkinson@bae.co.uk

Dr D. MEISTER

University of Kansas
Department of Psychology
1111, Wilbur Avenue
San Diego, CA 92109 205, USA
Tel: +1 619 270 9653

PANEL EXECUTIVE

From Europe and Canada:

Dr C. WIENTJES
RTA/NATO/HFM
7, Rue Ancelle
BP 25
92201 Neuilly sur Seine Cedex
France

From USA

RTA/NATO/HFM
PSC 116
APO AE 09777

Tel.: +33 (0)1 55 61 22 60
Telefax: +33 (0)1 55 61 22 99/98
Email: wientjesc@rta.nato.int

Technical Evaluation Report

by

Tom Kontogiannis, PhD

Department of Production Engineering and Management
Technical University of Crete
University Campus
Chania, Crete GR 73100
Greece

1. INTRODUCTION

The Human Factors and Medicine Panel (HFM) of the NATO Research and Technology Organisation (RTO, a merger of the former NATO Advisory Group for Aerospace Research and Development - AGARD- and the NATO Defense Research Group-DRG) held a workshop on "The Human Factor in System Reliability: Is Human Performance Predictable?" at the University of Siena, Certosa di Pontignano, Siena, Italy, 1-2 December 1999. The workshop was organised by Dr. David Embrey of Human Reliability Associates Ltd. as Chairman and Ms Jo Davies of ESE Associates Ltd. as Coordinator. The host was Dr. Antonio Rizzo of the University of Siena. The workshop audience included experts mainly from NATO countries. Eleven papers, including two keynote addresses, were presented from five NATO countries (Italy, Netherlands, Sweden, United Kingdom, and United States of America).

2. THEME

A fundamental part of the system design process involves the evaluation of the sources of potential human errors, their impact upon the successful operation of the system and potential methods for recovering errors or mitigating their consequences. Within the design of complex military systems, there is an increasing requirement to justify their reliability, safety and dependability by the application of techniques of formal risk analysis. A basic requirement of these techniques is the ability to predict the ways in which the hardware, human and software components of the system can fail and the consequences of failures. This allows the designer to choose a range of strategies, which may differ in cost, to minimise the probability of system failures. In order to perform comprehensive and cost-effectiveness analyses, there is also a requirement to quantify the likelihood of the potential failures revealed by the qualitative analyses.

Over the years, there has been considerable interest in both the qualitative and quantitative aspects of human reliability analysis from the designers of safety-critical systems in areas such as, nuclear power, transport, chemical processing, aviation and military systems. The application of these approaches has been limited by the unavailability of effective techniques for predicting human errors and the lack of reliable sources of data on human performance. The approaches to generating these data have tended to assume that human error data can be treated in the same way as that collected for hardware components. However, there are good reasons for believing that this is not the case. New approaches to data collection and human error quantification are needed which would take into account the unique characteristics of human operators. Papers were solicited that addressed the cognitive processes mediating the impact of workplace conditions on human performance as well as the wider organisational context that breeds human errors.

3. PURPOSE AND SCOPE

The workshop expected to review state-of-the-art knowledge about the following areas:

- Review of the fundamental differences between hardware, software and human performance and their implications for predicting human performance.
- Evaluation of the state-of-the-art of human factors knowledge with regard to its application to risk assessment studies in real-world domains.

- Review of techniques and practical tools for assessing human reliability and its dependence on workplace and organizational factors.
- Assessment of ways in which unobservable aspects of human performance (e.g. cognitive errors) should be treated, and the implications for data collection.
- Assessment of the extent to which the organisational factors underlying human errors need to be considered in military systems.
- Theoretical approaches pointing to new directions in carrying out research in human reliability in an applied military or industrial domain.

The workshop formed a link between two NATO working groups: the Research Study Group 25 (RSG25), which is in its final year and focused on data collection aspects of accidents and incidents, and the Working Group 30 (WG30) which is in its first year and aims to investigate the role of human reliability assessment techniques within the overall system design process.

4. WORKSHOP PROGRAM

The workshop was opened by Ms Jo Davies who introduced the audience to the theme of the workshop. Dr. Antonio Rizzo who also acted as a Local Coordinator also welcomed speakers and participants.

An overview of the general objectives of the RTO (after the merging of AGARD and DRG) and the HFM mission, scope and mode of operation was given by Dr. Cornelis Wientjes, the Executive of the HFM Panel. Following this, Dr. Wiel Jansen of TNO Netherlands presented the work carried out by the Research Study Group 25 while Dr. David Embrey presented the objectives of Working Group 30 (WG30) .

The papers were arranged to address 5 specific topic areas. Two keynote addresses were given at the start of each day covering broader issues:

Keynote address I: "Anticipating failures: What should predictions be about?" by Erik Hollnagel, University of Linköping, SE.

Keynote address II: "Impact of organisational factors on effective human reliability assessment" by James Reason, University of Manchester, UK

Session I was chaired by Peter Wilkinson, BAE Systems UK and addressed the specific question as to whether Human Performance can be addressed within the current Safety Assessment process. Contributions were received from Maarten Boasson, Signaalapparaten, NL and Ed Ridge, BAE Systems, UK who gave an impromptu overview of the Eurofighter Safety Assessment process.

Session 2 was chaired by Gretchen Burrett, Gregory-Harland, UK and addressed the qualitative and quantitative aspects of predicting Human Reliability. Contributions were received by Peter Wright, University of York, UK and Colin Drury, University of Buffalo.

Session 3 was chaired by Reiner Onken, University of Bundeswehr, GE and discussed the cognitive aspects associated with Human Reliability Assessments. Contributions were received from Neville Moray, University of Surrey, UK and Peter Roelofsma, Free University of Amsterdam, NL

Session 4 was chaired by Wolf Kaeppler, FGAN, GE and addressed the data collection aspects. Contributions were provided by John Chappelow, DERA CHS, UK and David Embrey, Human Reliability Associates, UK.

Session 5 was chaired by David Embrey, HRA, UK and addressed the organizational dimensions of Human Reliability. Contributions were received from Patrick Hudson, University of Leiden, NL and Antonnio Rizzo, University of Siena, IT

5. TECHNICAL EVALUATION

5.1 Keynote Address I

In his keynote address, (paper #KN1) Hollnagel gave a global view of past and current models of accident causation and examined their relationships to predictive models of human reliability. Changes in the conceptualization of human error over the last few years have been reflected in new developments in both retrospective and predictive analyses of human factors in system reliability. Classical ergonomics and error psychology have tended to view human error as a failure of the information processing system in cases where job demands exceeded human capabilities. This tradition generated practical models of accident causation concerned with the investigation of error mechanisms and the work conditions that triggered these behaviours. However, the relative sophistication of accident models has not been matched by failure prediction models. Unfortunately the direction of links between errors and causal conditions in *post hoc* analysis cannot easily be reversed in making error predictions. An increase in job demands, for instance, may not necessarily lead to errors since humans may compensate by changing their control strategy, or may rely on team communications for error detection, or make use of available system barriers. These adaptability, recovery and barrier functions have been the focus of current human reliability approaches that come under the framework of cognitive systems engineering.

Instead of focusing on human failures and error mechanisms, this new approach advocates that analysts should examine how working conditions combine together and influence human behaviour. In this sense, predictions should be more about working conditions and their influence than on failures and error tendencies. Hollnagel argued that analysts should pay particular attention to the interaction between 'context' (i.e., common work factors and system barriers) and 'control' (i.e., modes of performance and shifts when demands change). This is the underlying view of the CREAM technique (Cognitive Reliability Analysis Method) presented in the second half of this presentation. The concept of context has driven the development of a model of common performance conditions (e.g., available time, number of goals, communication efficiency) and a taxonomy of barrier functions (e.g., interlocks, work permits, instructions). On the other hand, modes of control can range from opportunistic behaviours to tactical and strategic ones. Although CREAM has already been used in *post hoc* analysis, its strength lies in making predictions about the interaction of context and control. This interaction should drive the calculation of human failure probabilities. Hollnagel concluded with a number of future research needs concerning theoretical and empirical studies of how performance conditions could affect the likelihood of losing control, studies of how barriers can fail, and requirements in terms of methods and data collection.

The following papers discussed issues concerning the role of cognitive factors in system reliability and presented several behavioural and analytical methods for predicting human reliability.

Boasson (paper # 1) discussed several difficulties and problems in addressing human error within current safety assessment processes. His main argument was that only routine aspects of human performance can be considered with traditional quantification methods while new approaches are needed to address "intelligent" behaviours such as, decision making and problem solving in the face of novel events. Within the current state of assessment methods, Boasson argued that the best that can be done would be to specify what constitutes acceptable operator performance under a wide variety of normal and abnormal process conditions. This repertoire of operator tasks and skills could provide the basis for designing operator interfaces and expert systems that would prevent the system going outside its safety boundaries. For instance, limiting functions could reject erroneous human actions and critiquing expert systems could detect human errors and provide appropriate explanations to operators. This error mitigation approach can be supplemented with other preventive approaches such as operating procedures and training regimes. A thorough specification of acceptable behaviours, therefore, would provide input to the design of operating procedures and training and foster conformance to the desired standards of performance. Both mitigation and prevention approaches, however, may face new challenges as systems become technologically more complex and compact. Critiquing expert systems, for instance, may fail to recognize erroneous performance in novel situations while operating procedures may restrict creative behaviours. These challenges to reliability during system operation should be addressed by new developments in the area of human reliability assessment. Boasson also argued that system

safety should integrate issues of system operation with issues pertaining to system design and system implementation. Therefore, safety and reliability issues should be addressed within the context of design, implementation and operation.

Moray (paper # 4) addressed the issue of how cognitive processes mediate the impact of work conditions upon human reliability. The human factors and ergonomics literature has quite a strong armamentarium of quantitative models of human performance (e.g., models of signal detection and control of attention). While these models appear to provide valuable data for a range of skill-based and rule-based tasks (e.g., scanning instruments, inspecting equipment, following instructions and manual tracking) they are limited in studying cognitive or knowledge-based tasks. In order to understand how operators engage in cognitive tasks (e.g., fault diagnosis and problem solving) Moray advanced the concept of 'mental models' in mediating the perception-action cycle of performance. A mental model is a 'knowledge structure' or a 'cognitive map of the world and its possibilities' that can help operators adapt to variations in the work environment. For instance, knowledge of system dynamics may direct eye-movements and increase sensitivity to particular aspects and interpretations of available information. Unavoidably, these perceptual processes will sometimes uncover data that the mental model does not expect or fail to find data that it does expect; thus, mental models can be updated or modified and become calibrated to the characteristics of the complex environment. The functioning of mental models and their interactions with the processes of perception and attention is of paramount importance to understanding how people manage multiple tasks in difficult situations. Moray proposed that scheduling theory may be a good candidate for a unifying framework in the study of strategic aspects of behaviour. If we consider cognitive functions as resources and the objects of those functions (namely, cognitive tasks) as jobs, then we should be able to benefit from the work done in a number of engineering disciplines where scheduling theory has been applied for many years. In a sense, mental models are useful in developing representations of tasks, their demands and priorities; scheduling theories, on the other hand, can be valuable in understanding strategic aspects of performance (e.g., queuing of interrupted or upcoming tasks and allocation of tasks to cognitive functions). Therefore, more research is needed into these cognitive functions in order to develop quantitative models of human performance that would predict error modes and underlying causes.

Roelofsma (paper # 5) presented a study on human performance which demonstrated the benefits of the experimental approach over more analytical methods of system evaluation (e.g., task analysis and error checklists). An experimental approach to system evaluation would simulate user interactions with the system, at some level of fidelity, and test human performance over a range of tasks. This is a more laborious effort than analyzing user interactions, making error predictions and finally, assessing the overall system reliability. However, many insights can be obtained from experimental studies with regard to how users adapt their performance when job demands change. A decrease in memory demands, for instance, may not necessarily give rise to superior performance since user behaviour and strategy may change as well. Roelofsma carried out an experiment to test the effect of memory-aiding upon human performance in a command and control task. A simulation was developed where subjects were required to make trading decisions in a business environment by buying and selling commodities to trading centers. Two experimental groups started with the provision of a memory aid which was subsequently removed in one of the groups. Two other groups started without any memory support but one of them allowed access to the aid at a later stage. Decision making performance was measured in terms of a success score (i.e., profit making), a failure score (i.e., bankruptcies), decision, speed and information search profile. In general, the results showed that memory-aiding did not affect the mean success score for each decision or the overall failure score. On the contrary, memory-aiding reduced the amount of searching for new information. It appeared that the availability of the aid prompted subjects to spend more time in processing the available information at the expense of monitoring event changes in the dynamic environment. The most plausible explanation for the results related to the adaptation in performance when removing or introducing the aid. Specifically, the introduction of the aid prompted subjects to adopt an analytical decision strategy, spending more time in evaluating alternative options than searching for changes in the environment. On the other hand, lack or removal of the aid tended to reinforce a more intuitive strategy whereby alternatives were evaluated in a sequential fashion; this enabled subjects to spend more time in searching for new events and assimilating more information from the environment. Under these conditions (i.e., low expertise and high uncertainty), the type of memory-aiding chosen was ineffective. It is conceivable that other forms of memory-aids could make a better impact in performance, especially when

the preferred strategies of participants are taken into account. The implication is that experimental tests may provide a good basis for evaluating changes in man-machine systems. However, the issues of simulation fidelity, task type, and individual differences should be taken into account when deciding on aspects of system reliability that should be explored experimentally or analytically.

Wright (paper # 2) reported on an analytical approach how to perform qualitative assessments at the early stages of system design. He emphasized that quantitative assessments can be used at a later stage to examine the extent that a system conforms to a set of usability criteria. The qualitative approach uses a technique for human error prediction, known as THEA, which provides feedback to an iterative design process. THEA uses the concept of 'work scenario' to reflect on current thinking about the role of work context on human performance. A work scenario is a thorough description of agents and their responsibilities, the task carried out, the procedures used, the environment in which the activity takes place, and the history of tasks (e.g., successful and incomplete tasks) that led to the current system state. Two other important elements of the work scenario include the technology or tools used to perform the tasks and the exceptional circumstances associated to the scenario due to variations in agents, situations and tasks. A variety of data sources should be used to specify the work scenario including, experience with earlier versions of the system, incident reports, and changes in technology (e.g., two versus three pilots in the flightdeck). The phase of scenario generation is followed by the identification of human errors. To this extent, Wright advocated the use of behavioural and cognitive error checklists, the later referring to the cognitive aspects of performance that give rise to certain behavioural acts. The cognitive error analysis, however, has been based on models of cognition proposed in earlier years by Donald Norman and Jens Rasmussen. THEA appears to be a promising analytical methodology in the sense that it supports system designers to analyze the whole context of work and identify opportunities for preventing errors, enhancing recovery, and mitigating error consequences. This formative assessment provides early feedback to inform system designers. At a later stage, quantitative assessments can be made by using HEART, an existing technique of assessing failure probabilities. The task conditions and their relative importance identified earlier can provide input to HEART in order to generate error probabilities. However, as Wright emphasizes, the primary objective of using THEA at this later stage would be to make comparisons between different design features rather than obtain conditional probabilities for risk analysis. In other words, error probabilities are not treated as objective truths but rather as starting points for discussion.

The presentation of Drury (paper # 3) demonstrated how good human factors knowledge, in terms of performance models, can be used to combine an analytical and behavioural approach to quantification. Human reliability in aircraft inspection tasks is very important for setting up proper inspection intervals; too few inspections may give rise to accidents whilst too many can increase costs. Drury has reviewed human factors studies on non-destructive testing, industrial inspection and maintenance resource management in order to develop a quantitative model of the aircraft inspection process. Such a model would examine the stages involved in inspection (e.g., search and decision), the variability of performance in inspecting different faults (e.g., cracks, deformation, corrosion) and the impact of contextual factors. Although the existing literature provided useful insights, Drury identified a number of limitations; non-destructive inspection, for instance, focuses on one defect type and on one dimension whilst industrial inspection lacks face validity. He developed a five-stage model of inspection (initiation, access, search, decision, response) and identified the factors that affect two of the more error prone stages, that is search and decision. Peripheral visual acuity, for instance, affects fixation area and, thus search, whilst the cost of a miss or false rejection affects the decision stage. To furnish this model of aircraft inspection, Drury performed a series of experimental studies. In his paper, there is a succinct description of the Visual Inspection Research Program (VIRP) undertaken for the FAA where a retired Boeing 737 test aircraft was used. Twelve experienced inspectors performed ten tasks under highly realistic conditions in a flight hangar. The results showed that inspectors took 7.5 to 12.3 hours for the ten tasks. On a set of large cracks and corrosion defects, which the manufacturers would expect inspectors to find, the probability of detection was also quite variable ranging from 0.5 to 1.0 on large cracks and from 0.3 to 0.6 on large corrosion areas. There was little evidence of a speed/accuracy tradeoff across inspectors. There was also low correlation between inspector performance on the 10 tasks as well as between pre-test measures and task performance. A more detailed analysis was undertaken for one task which was video-taped in order to identify search and decision errors. Search performance could be characterized as consistently poor, whereas decision performance was better, but highly variable. Search and decision performance were statistically

unrelated. Such findings allow us to focus interventions, for example by improving lighting and training to support search, or by using training and feedback to reduce inter-inspector variability in decision. Forthcoming studies will examine how the ‘probability-of-detection’ curve is affected by different types of defect as well as by different conditions of work.

5.2 Keynote Address II

In his keynote address (paper #KN2), Reason has shifted the focus of discussion from cognitive factors to organisational factors and the workplace culture. Over the last few years there has been an increasing recognition of the impact of organisational factors upon system reliability. There has also been an awareness that system safety has two faces, namely ‘occasional vulnerability;’ and ‘resilience’. While human error has been implicated in some 70-80% of bad outcomes, the human operator continues to protect the system in a dynamic and uncertain world. Reason has pointed to a paradox in the variability of human performance. On the one hand, elimination of human error has been seen as a primary goal by many managers; as a result, organisations strive for greater consistency of human action (e.g., through procedures). On the other hand, human variability has been quoted as a major source of system protection (e.g., through innovation) in various incidents including Apollo13, Davis Bessie, Gimli Glider and United 232. Hence, ensuring effective compensation, error recovery, and improvisation would call for a special kind of organisational practices and culture; these factors have permeated the concept of the high reliability organisation (HRO). Reason argued that current approaches have focused on event-dependent analyses of human performance which far outweigh event-independent observations. He quoted Weick arguing that human reliability should be seen as a "dynamic non-event" and this is best studied by continuous observation. It is a "non-event" because most of the time nothing happens as operators are able to compensate; it is also "dynamic" because safe outcomes (non-events) are achieved through timely adaptations of human operators to an uncertain and dynamic world. Drawing upon Weick, Reason proposed that high reliability organisations exhibit five main characteristics, that is, (i) a continuing awareness of the possibility of failure, (ii) an expectation that errors will be made but trained personnel should be able to recover them, (iii) a reporting culture regarding near misses and incidents, (iv) a generalized rather than localized approach to failure identification, and (v) a contingency planning practice whereby failure scenarios are anticipated and coping plans are thought of in advance. Further research on the variability aspects of performance and the impact of organisational practices has been undertaken by Reason in a study of neonatal switch operations performed by cardiac surgeons. Data were collected on 230 surgical procedures performed by 21 UK surgeons whilst detailed observations were made on 165 cases. The results showed that failure rates in these subtle operations were 6.5% for deaths and 18.5% for near misses. Surgeons were able to compensate for almost half of major events and 80% of minor events that occurred during these operations. Observations indicated that good compensators were wary of possible contingencies and mentally rehearsed ways of coping with them ahead of time. ‘Intelligent wariness’ and ‘preparedness’ were the key elements of effective compensations. This study demonstrated that field observations could be very valuable in generating quantitative data about human error as well as about error recovery. The point has also been made that the time has come for looking deeper into the practices of high reliability organisations that increase systemic resilience to hazards and nasty surprises.

The papers on the second day have drawn upon this framework that views system reliability within the wider organisational context. Elaborations on the role of safety culture have been followed by practical techniques for quantifying the influence of workplace and organisational factors, requirements for data collection, and methods for capturing safety knowledge.

Hudson (paper # 8) elaborated on the issue of safety culture and presented a systemic approach for high reliability organisations. Drawing upon and extending the work of Westrum, he proposed that corporate cultures can range from pathological (i.e., whereby safety practices are at the barest industry minimum) to generative ones (i.e., whereby all employees participate and share responsibility for safety). Making progress towards safety can be seen as going through a number of intermediate steps and ultimately achieving a generative culture. Between the two extremes there are another three levels, that is (i) reactive cultures (i.e., keeping just one step ahead of regulators but showing concern about accident trends), (ii) calculative cultures (i.e., calculating the odds based on what went wrong last time but failing to appreciate human factors), and (iii) proactive cultures (i.e., recognizing the importance of organisational factors and getting ahead of

problems). Hudson argued that organisations can be placed at some point along this continuum and that safety culture has to evolve; steps cannot be skipped to the generative culture. A model has been presented to understand how beliefs and attitudes can influence organisational behaviour and how barriers can get on the way to implementing the desired safety practices. Hudson suggested a taxonomy of organisational attitudes (i.e., termed the "talk" factor) and a taxonomy of organisational behaviours (i.e., termed the "walk" factor) that could guide interventions in safety culture; in fact, the walk/talk ratio could be seen as a measure of development. Examples of organisational behaviours may include: dealing with change, reaction to trouble, risk appreciation, safety procedures, rewards for good performance, and level of care. It is also very important to understand the barriers to this process and the counter-pressures that may force organisations back to a calculative culture. Hudson perceived of an 'addiction model' that can block organisational changes in safety culture and pointed to certain ways of overcoming addiction. There is a need to understand the context and dynamics of change since organisations can even regress from the generative stage. Environmental factors (e.g., a less advanced culture of the regulatory authorities) and internal factors can hold developments back. Learning from the past, adapting the organisational structure to the tempo of the situation, and maintaining 'intelligent' wariness require continuous effort and commitment. Further research into safety cultures is a way forward in enhancing system reliability.

An approach that aims to provide the link between human performance models, direct and indirect or organisational factors in accident causation has been taken by Embrey (paper # 7) in the presentation of the Contextual Causal Model (COCAM). Embrey argued that existing causal models of human error are based on generic models of human performance and make it difficult to incorporate end-user knowledge of factors known to influence error in a specific domain. A contextual approach to error causation would focus on the performance mechanisms pertaining to a specific context, incorporate contextual knowledge held by end-users, and take into account the wider organisational context (e.g., procedures policies, training, safety culture). These objectives have driven the development of the COCAM model which has found extensive application in several industrial domains – e.g., rail transport, marine industry, nuclear power operations, and aircraft maintenance. Influence diagrams are used as graphical methods for representing the causes of human error and system failure at different levels (e.g., performance mechanisms, workplace factors and organisational policies). An iterative process is followed in building the COCAM model of an event whereby several incident and near miss reports are reviewed in conjunction with available research in the specific domain. A preliminary influence diagram is drawn which is modified as more knowledge accumulates by interviewing end-users, designers, and line supervisors. This process results in an influence diagram of direct and indirect causal factors that are evaluated in terms of their relative contribution to the final event. These weights of importance are used in combination with ratings of the quality of these factors in order to generate an overall index of failure or probability that the final event will occur. A software tool has been developed to assist analysts in error quantification. Cost-effectiveness analysis is also possible in this software by assessing how changes in the quality of a causal factor can affect the event probability and by assigning costs in implementing such changes. Embrey demonstrated the COCAM model in an assessment study of train drivers passing signals at danger. Performance mechanisms - such as, signal visibility, attention focus and alertness - are initially identified for different stages of human performance. The analyst can extend this level of description by re-describing each performance stage (e.g., attention focus) in terms of other context-specific mechanisms - such as multiple-tasking, signal position cues, route knowledge, and distractions. The influence diagram continues with causal factors at the level of the workplace (e.g., weather conditions, obstructions, design of signal devices) and the organisational level (e.g., maintenance policy, route training policy, procedures policy). Generic weights of importance can be produced by aggregating data from several incident reports and available research which can be subsequently modified as more knowledge is gathered. In this sense, the analysis of previous incidents and near misses provides valuable input to the prediction of causal factors and human errors that can lead to the final event (i.e., signals passed at danger). Improvements in the methodology will be forthcoming as human performance models become more elaborate.

The issue of human reliability assessment is particularly important in military aviation where fast response systems are in operation. Chappelow (paper # 6) has presented a current project undertaken by the Defense Evaluation and Research Agency (DERA, UK) to develop an incident coding system. The accident database should capture crucial features of causal factors and provide input to risk analysis. By using historical data to estimate the quality of underlying causal factors and the strength of their influence on error mechanisms,

relatively objective sensitivity analysis would be made possible. The need for a classification scheme of different types of incidents had long been recognised by aviation psychologists and ergonomists. Task taxonomies for certain perceptual-motor tasks, performed by pilots, have been developed and proved very useful in extrapolating reaction times across several types of emergencies. However, other pilot tasks demanding more interpretation or complex decision-making have challenged existing databases and required more elaborate taxonomies of human errors and causal factors. A fact not evident in earlier accident analyses and databases was the strong influential character of social factors in military aircraft accidents. A recent review of social factors in accidents by Chappelow and O' Connor identified not only communication problems and decision-making biases but also organisationally induced tendencies to more risky behaviour. Chappelow has sought to develop a causal factors database in a way that human error and machine failures could be described in compatible terms. His taxonomy of causal factors was cast at different levels including, environmental factors, enabling factors (e.g., ergonomics and training) and predispositions (e.g., personality, fatigue, overarousal). An influence diagram approach was that was similar to the COCAM model. Chappelow found that the influence diagrams generated by the database were much more elaborate than the ones produced by teams of experts. In order to obtain reliable estimates of error rates there is a need for focused efforts on the creation of an open reporting culture. Chappelow quoted a study of collecting data on error rates in designing and using seat ejection pins where the reported near miss cases were a magnitude of two greater than the reported accidents. This brings into fore the earlier discussion on the role of safety culture in system reliability. Some of the goals of operators are determined by the design of the system while others are influenced by the teams they work in and the organisation as a whole. In addressing system reliability, thus, we need to consider not just the man-artifact system but also the whole organisational context in which artifacts are used.

One of the most important aspects of system safety, as advocated by the contextual approach, is safety knowledge concerning the use of organisational resources (e.g., humans and artifacts). Rizzo (paper # 9) presented a proactive method, SHELFS, for capturing safety knowledge in organisations that do not have a long tradition in ergonomics. He emphasised that safety knowledge is not only about human errors and equipment failures but also about safe working practices and other vital signs of safety. This is a proactive approach consonant with the view of safety as a "dynamic non-event". Rizzo has built on the SHEL model of Edwards (Software, Hardware, Environment and Liveware) in order to develop a new method for capturing safety knowledge. The SHEL model has been enriched by incorporating the 'cultural-historical' framework of Vygotsky and the 'distributed cognition' approach of Norman. The 'cultural-historical' approach views 'knowledge' as embedded in the interactions between users and artifacts; hence, the evolution of artifacts over time and their differences with other similar artifacts conveys important information about work practices. The tradition of 'distributed cognition' addresses the social interactions surrounding artifact use and resource allocation; hence, the mapping of artifact and human resources is an important aspect of how safety issues are managed. Rizzo has introduced the new SHELFS model to the Italian National Railways (FS) by selecting and training certain operators (called Line Tutors) whose role was to identify critical safety issues and propose adequate solutions. The SHELFS method involves three phases whereby the work process is described in terms of a matrix workflow, critical issues are identified, and solutions are proposed on the basis of several meetings with representatives of all parties involved in safety. The matrix flow (first phase) aims to map the main classes of resources involved in the technical process. It represents the process in terms of its basic activities, the personnel involved, the communication flows, the procedures and rules involved and the hardware elements in use. This provides the required input to the second phase where the Line Tutor investigates the real breakdowns experienced by workers in performing these processes and the related causes. A hierarchical taxonomy of performance breakdowns has been developed concerning the hardware, software and liveware aspects of the SHELFS model. The aim of this phase is to examine, according to operational experience, how well the resources interact in the existing organisational context. The discussions with end-users required in the second phase of work are also valuable in collating knowledge about ways to overcome poor mappings between people and artifacts. This information is used in the third phase of SHELFS where the Line Tutor holds several meetings with the representatives of all human roles necessary to carry out the process at hand. In this phase of work, solutions are evaluated and the most effective ones are selected. An application of the SHELFS methods in the area of train maintenance activities showed that a great deal of safety knowledge came into fore which was previously embedded in the day-to-day activities, but unknown to the safety department.

6. CONCLUSIONS AND RECOMMENDATIONS

The presentations and discussions held in this workshop have provided useful insights into current philosophies and methodologies in assessing the role of the human factor in system reliability. It became evident that new approaches to human reliability should focus more upon the cognitive processes and the organisational context within which behaviour takes place. Papers in the first three sessions focused on the cognitive processes mediating the effect of work conditions on human reliability. It is important to understand how human operators adapt their control strategies and behaviours to changes in the demands of the situation before we are able to quantify human reliability. Two of the most prevalent traditions in cognitive ergonomic and psychology model cognitive processes in terms of mental models and strategic modes of control (e.g., ranging from opportunistic to strategic control). Being reliable in operating a complex system, therefore, entails a state of alertness and preparedness in updating one's own mental model of the problem and changing to the most appropriate mode of control. The current state of human factors knowledge that we have about these human adaptations enables us to apply fairly robust models of human performance to tasks that appear to be routine and familiar to the operators. We also appear to get a good grasp of the cognitive process that we should consider when addressing the more cognitive tasks (e.g., fault diagnosis and problem solving) but we still lack proper methodologies and models of performance.

Some of the papers have presented new tools for assessing human reliability. The classical concept of a human task has been replaced by the concept of 'work scenario' (e.g., in the THEA method) which encapsulates the user interactions with the system in order to perform a specific task; this approach takes into account the influence of contextual factors on human reliability. The interaction between work context and cognitive control has received a lot of attention in the CREAM technique which has recently been furnished with a taxonomy of barrier functions. Both methods, however, are more concerned with a rank ordering of the criticality of tasks rather than precise quantification of failure rates. It appears that analytical approaches should be combined with field studies in order to be in a better position to assess reliability in a quantitative fashion. The study of inspection reliability in aircraft maintenance illustrates this need for a combined approach.

The second day of the workshop has been more concerned with the organization context of system reliability. The presentations of Reason and Hudson have provided a fundamental framework for addressing organisational factors and safety cultures in system safety. Both researchers have been involved in studying the essential features of high reliability organisations. Awareness of failures, error recovery training, a reporting culture, and a contingency planning approach are essential ingredients of high reliability organisations. What has become apparent by the presentations of Reason and Hudson is that operators and organisations need to be in a state of 'intelligent wariness' where worst case scenarios are mentally rehearsed and contingency plans are formulated ahead of time. This sort of compensatory behavior was characteristic of cardiac surgeon who managed to recover from several events during neonatal operations. The same behavior is also required even of generative cultures since the danger of regressing back to calculative cultures is ever present.

A method that seems to take into account the complex ways in which workplace and organizational factors interact is the COCAM method proposed by Embrey. The influence diagram approach to the assessment of human reliability appears very promising in that the knowledge of end-users is integrated with the human factors knowledge. At its current state of development, COCAM relies to some extent on expert judgments about the relative importance of causal factors. In this sense, it is a useful approach for organizing existing human factors knowledge. Further developments in human factors and human reliability studies are needed in order to make less use of subjective judgments. A similar approach has been taken by Chappelow in designing a database for aircraft accidents. By using incident reports, near miss reports and questionnaires it was possible to gather quantitative data about human failures. These data can provide useful input to risk analysis.

A concept that has been put forward in this workshop is that reliability is a "non-event" phenomenon. This implies that analysts should be concerned not only with human errors but also with safe working practices and other vital signs of safety. This safety knowledge is an asset for high risk industries. The method proposed by Rizzo, SHELFS provides a practical way for putting into practice this new concept of system safety. It is also worth noting how the 'distributed cognition' approach and the 'cultural-historical' framework of activity

theories (e.g., Vygotsky) have been fed into the SHELFs methodology. These two approaches have also been implicated in the concept of 'work scenario' applied by Wright in his THEA methodology.

In summary, the workshop has brought into light new approaches to the assessment of system reliability. Most of them are concerned with the role of cognitive factors and organisational factors in safety and have been applied to a variety of industrial projects. It seems that our current models of human reliability have been enriched with more performance mechanisms and influential factors. However, there is a long way to go in order to specify the links between performance mechanisms and underlying factors. More field studies and simulation-based studies are needed in order to delineate these relationships so that we are in a better position to make quantitative predictions about human performance.

Anticipating Failures: What Should Predictions Be About?

Erik Hollnagel

Professor, Ph.D.

Graduate School for Human-Machine Interaction

University of Linköping, SE-581 83 Linköping

Sweden

eriho@ikp.liu.se; eriho@ida.liu.se

Summary: Accident analysis and performance predictions have traditionally been pursued in separate ways, using different concepts and methods. This has made it difficult to use the experiences from accident analysis in performance prediction. As a result, performance prediction is still focused on the concept of individual “errors”, despite overwhelming evidence that accidents are caused by a concatenation of conditions rather than a single action failure. It is argued that the anticipation of failures should be based on better models of how performance conditions determine actions, and that the inherent variability – or unreliability – of human performance is the noise rather than the signal.

1. INTRODUCTION

Accident analysis and performance prediction for human-machine systems have traditionally been pursued as two separate activities, despite the obvious fact that they refer to the same reality – namely the occurrence of unexpected events leading to unwanted outcomes. Accident analysis has been concerned about unravelling the complex of causes that might explain what happened, and preferably finding one or a few causes that could be considered the root or origin of the accident. Performance prediction has been concerned with trying to identify in advance the risks inherent in a system, in order to be able to change or modify the design so that these risk can be reduced or eliminated. In both cases a common motivation has been the dramatic rise since the 1970s in the number of cases where the causes of accidents have been attributed to incorrectly performed human actions. Although this does not by itself mean that there have been more “human errors”, it expresses a distinct change in attitude towards the analysis of accidents and the commonly accepted set of causes (cf. Hollnagel, 1993a).

Accident analysis for systems involving human-machine interaction has always had a strong psychological flavour, looking toward “human error mechanisms” and various deficiencies of information processing that are supposed to occur in the human mind (e.g. Senders & Moray, 1991). In contrast to that, performance prediction has been dominated by the engineering quest for quantification, as epitomised by the PSA event tree, and models and methods have been constrained by that (e.g. Dougherty & Fragola, 1988). In both cases there has been a strong predilection for considering “human error” as a category by itself, referring either to complex models of how information processing can go wrong or to estimates of single “human error probabilities”. This view persists despite a growing realisation that it is a gross oversimplification which fails to recognise the complexity and significance of human performance failures (Hollnagel, 1993a; Woods et al., 1994).

2. APPROACHES TO ACCIDENT ANALYSIS

The analysis of an accident is always based on an accident model, i.e., a conceptualisation of the nature of accidents, specifically how a set of causes and conditions may lead to an accident. Current accident models must account for the complex interaction between humans, technology, and organisations. The accident model may be explicitly formulated but is more often implicit, hidden in the assumptions that investigators make. Every accident model is based on the principle of causality, which states that there must be a cause for any observed event, and the models serve as guidance for finding the acceptable causes. In the following I will briefly consider the major changes to accident models since the 1950s, since these reflects the developments in the commonly agreed understanding of the nature of an accident.

2.1 Simple Accident Model

The first accident models tended to see accidents as caused either by failures of the technology or incorrect human actions, cf. Figure 1. Before the accident the system was assumed to be in a normal state, and an incorrect human action was seen as the primary cause of the accident. Accident classifications typically used the “human error” category as a kind of catchall, or garbage can, for accidents that could not be attributed to the failure of a technical component. The simple accident model corresponds to methods such as root cause analysis (Park, 1987; Cojazzi, 1993; Cojazzi & Pinola, 1994), which from a psychological view are relatively unsophisticated. In relation to the specific issue of human failures, the simple accident model is closely associated to the information processing point of view, which harbours three basic assumptions. Firstly, that there are reliable criteria of validity against which it is possible to measure a deviant response. Secondly, that psychological factors affect information processing and act to bias responses away from the standards considered appropriate. And finally that the human information processing system comprises a diverse range of limitations that are invoked under particular information processing conditions.

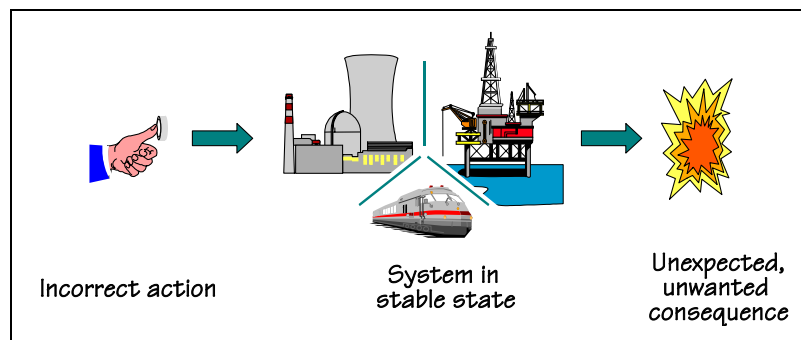


Figure 1: A simple accident model

2.2 Intermediate Accident Model

The simple accident model was gradually extended to recognise both the contribution of latent system states, and the complexity of conditions that could lead to an incorrectly performed human action, cf. Figure 2 – eventually ending by the extreme notion of “error forcing” conditions (Cooper et al., 1996). The complexity of working conditions relaxed the strong assumption of “human error mechanisms”, and encouraged descriptions of how human actions were affected by the conditions under which they took place. The latent system conditions – originally called latent system failures (Reason, 1992) – can be precarious conditions brought about by unsound practices of work, as well as consequences of earlier failures. As the name implies, the latent conditions remain undetected until changed circumstances turn them into manifest failures that require rapid responses – usually on top of other events that demand attention. Latent system conditions in safety functions are particularly malicious, because they decrease the safety level without anybody knowing about it while the process is running. In addition, when the safety system is needed, the lack of appropriate responses may lead to a temporary or permanent loss of control of the situation.

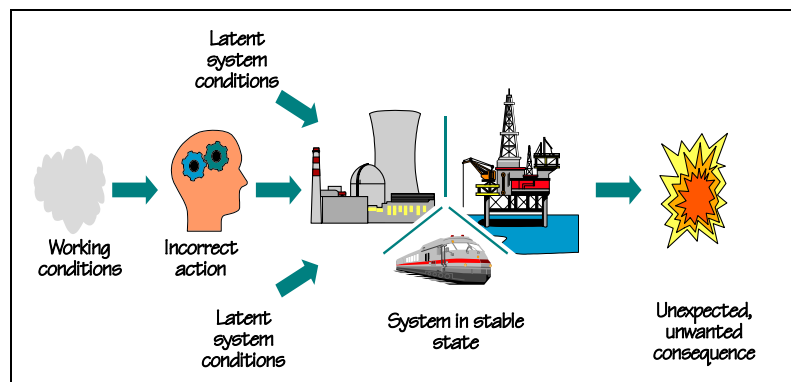


Figure 2: Intermediate accident model

2.3 Contemporary Accident Model

The common approach to analysing and understanding accidents has in the 1990s further shifted the perspective from individual actors to the organisational context. Although the actions – and failures – of individuals still constitute the initiating event, it is necessary to understand the complexity of the working environment, not least the existence of latent conditions. An excellent account of this work has been provided by Reason (1997), which emphasises the concept of organisational safety and how defences may fail.

In the current approach, as shown in Figure 3, the immediate or proximal cause of the accident is a failure of people at the sharp end who are directly involved in the regulation of the process or in the interaction with the technology (Reason, 1990; Woods et al., 1994). A combination of factors that relate to either the human, the technological or the organisational parts of the system – the so-called Man-Technology-Organisation or MTO perspective – is used to explain this failure. The failure at the sharp end is, however, only the triggering condition. The accident does not occur unless there is also a number of latent conditions that suddenly become “active”. Furthermore, the outcomes of the failure at the sharp end are both overt and hidden consequences, the latter possibly becoming latent conditions that during a future event may affect the safety of the system.

In addition to the immediate cause, this view also assumes a set of background or proximal causes that are due to function failures at the blunt end. People at the blunt end are to a large extent responsible for the conditions to which by people at the sharp end are exposed, but are themselves isolated from the actual operation. They can be managers, designers, regulators, analysts, system architects, instrument providers, etc. It is the ambition of the contemporary perspective to account for the complex interactions of distal and proximal causes, as well as for the temporal relations, i.e., the way in which past, present, and future are coupled.

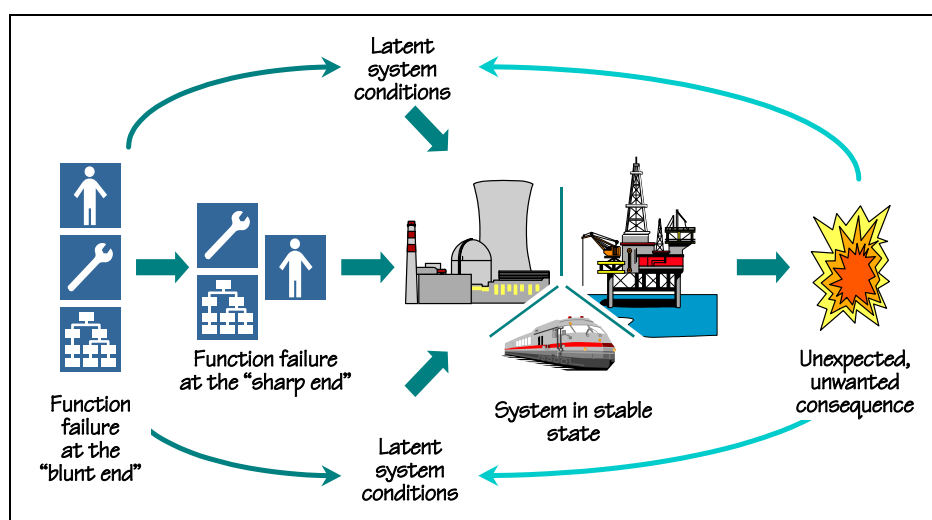


Figure 3: A contemporary accident model

2.4 The Nature Of Causes

Despite these developments, specifically the increasing sophistication in accounting for the organisational determinants of accidents, there is an almost intransigent preference to refer to “human error” as a singular concept. This preference persists in spite of the clear demonstration from the history of accident analysis that the notion of a cause itself is an oversimplification. As pointed out by Woods et al. (1994), a cause is an attribution after the fact or a judgement in hindsight, rather than an objective, unequivocal fact. The determination of the “cause” is a relative rather than absolute process, hence pragmatic and social rather than scientific and deductive. According to this view, a cause can be defined as **the identification, after the fact, of a limited set of aspects of the situation that are seen as the necessary and sufficient conditions for the effect(s) to have occurred**. A cause is in general acceptable:

- If it can unequivocally be associated with a system structure or function (people, components, procedures, etc.).

- If it is possible to do something to reduce or eliminate the cause within accepted limits of cost and time.
- If it conforms to the current “norms” for explanations.

This acknowledgement notwithstanding, accident models are firmly entrenched both in the idea that a “true” or root cause can be found, and in the idea that “human errors” necessarily must be part of the explanations. The result is that accident models become oversimplified, as shown by the left side of Figure 4. According to this view, the accident is first characterised in terms of the external error mode. Next, a suitable cause is expressed as a combination of likely psychological “error mechanisms” and performance shaping factors, where the latter can only exert their influence through the former. The contrasting view, shown by the right side of Figure 4, is consistent with the principles of cognitive systems engineering (Hollnagel, 1998a; Woods, et al., 1994). The search for causes necessarily begins in the same manner by the external error mode or manifestation of the performance failure. But rather than assuming that the proximal cause must necessarily involve a “human error mechanism”, or indeed even be attributable to an individual, the search considers the context of the socio-technical system as a whole. In the remaining part of this paper I will argue that this difference in perspectives has significant consequences for how performance predictions are made and how failures are anticipated.

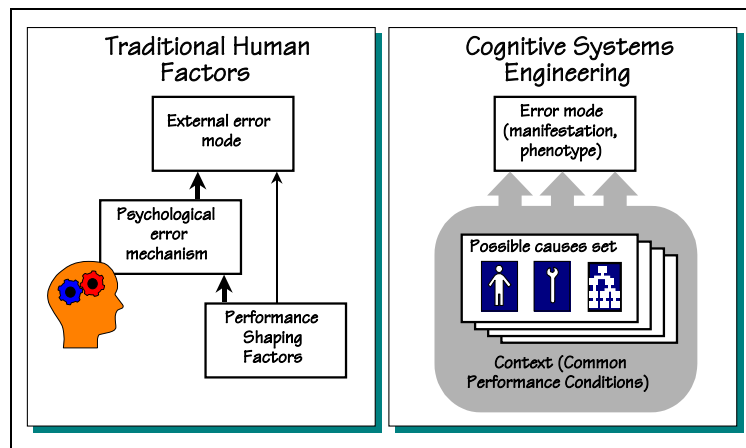


Figure 4: Two perspectives on causation

3. APPROACHES TO PERFORMANCE PREDICTION

As mentioned in the introduction, performance prediction has on the whole been separated from accident analysis. The reality of this separation becomes obvious if one tries to apply any of the established “human error models” for prediction. Indeed, neither the methods nor the categories used allow an easy reversal of the direction of going into the past to going into the future.

One reason for this is that accident models focus on **error types**, where as performance prediction must focus on **error modes**. An error type is a category that is based on and derives its meaning from an underlying model of human action – specifically of human information processing or “cognition in the mind”. Well-known examples of model-defined error types range from errors of omission and commission to skill-based, rule-based, and knowledge-based lapses and mistakes. An error type is linked to a specific model of the processes behind human action, and to how these processes may fail. In contrast to that, an error mode – or human failure mode – is a category that refers to a description of observable manifestations. Error modes may even be logically defined by referring to the small number of physically possible failures (Hollnagel, 1993b). Error modes thus refer to aspects of duration, time, direction, force, sequence, etc. As an example, performing an action too late is an error mode that refers to the aspect of timing of an action. “Action too late” is an overt manifestation and does not in itself make any assumptions of what lay behind it. Performing an action too late may also be described in the language of error types, i.e., referring to what the underlying cause was. Depending on the theoretical stance of the analyst, it may be described as an error of commission or as a rule-based mistake. For accident analysis it may be important to construct an acceptable explanation of the conditions and causes that lead to an accident, hence to focus on error types. For performance prediction it is

important to identify the types of incorrect actions that can occur, regardless of what the causes may be, hence to focus on error types.

Performance prediction has traditionally been pursued as a separate activity in the form of Human Reliability Assessment (HRA), which is the established way of finding the human failure probabilities required by Probabilistic Risk Assessment (Kirwan, 1994). Performance prediction is, however, also an integral part of system design. During the design process choices are made which involve assumptions about the responses of humans and technical systems in future situations. As commonly practised, system design has two major objectives. The first is to ensure that the system performs as required, i.e., that it meets the functional requirements. The second is to avoid that unexpected events happen and that failures occur. The former usually takes up the major part of the design process, whereas the latter is treated more sporadically – almost as a stepchild.

The concern for system failures has grown significantly over the last decades almost as a realisation that system failures generally are unavoidable (Perrow, 1984). The anticipation of system failures is guided by the dominating scientific paradigm, which traditionally is one of decomposition – in particular the decomposition of a system into its “natural” parts, humans and machines. This paradigm has been firmly established by disciplines such as human factors (ergonomics) and human-computer interaction. Since the reliability of modern technology is quite high, the logic of the decomposition approach has forced the focus onto issue of human reliability, usually as single individuals and more rarely as groups or organisations.

3.1 HRA And Human Performance Failure

Due to the influence of accident analysis and HRA, the common approaches to performance predictions have focused human performance – or rather, human performance failures. Performance prediction, as practised by HRA, confines itself to an investigation of the ways in which actions can possibly fail, often referred to as action error modes – or just error modes. In doing so, the likelihood of failure is seen as an attribute of human actions *per se*, often expressed in terms of a “human error probability” (HEP). This is quite consistent with the information processing view, where specific internal “error mechanisms” are assumed to exist. If a function can be seen as an attribute of a component, it follows that the possibility of function failure can be considered for the component by itself, although it is acknowledged that the circumstances or context may have some influence. In HRA the circumstances have been encapsulated by the set of performance shaping factors, which exert their influence in a simple, additive fashion. Yet the likelihood of a component function failure – read: “human error” – is calculated or assessed prior to, hence independent of, the effects of the performance shaping factors.

Anticipating failures of joint human-machine systems requires an underlying model. This should not be a model of human information processing in disguise, but a model of how human performance is determined by – hence reflects – the context or circumstances, i.e., a model of joint system performance rather than of human actions. This type of model corresponds to the notions of distributed or embedded cognition (Hutchins, 1995), although neither of these have been used to consider performance prediction specifically. A concrete expression of these ideas is found in the contextual control models (Hollnagel, 1998b), which describes how humans and technology function as joint systems, rather than how humans interact with machines. The contextual control models emphasise how human-machine co-operation maintains an equilibrium rather than how human-computer interaction can be optimised. The emphasis is thus on “cognition in the world” rather than “cognition in the mind”.

3.2 “Human Error” As Noise Or Signal

It is assumed both by HRA and accident analysis that it is reasonable to consider the inherent variability of human performance by itself, specifically that a performance failure is an attribute of the “human component” rather than of the circumstances during which actions take place. In this sense the “human error” is –

metaphorically, at least – the signal rather than the noise. This assumption is strangely inconsistent with one of the main tenets of the information processing approach, which states that:

“A man, viewed as a behaving system, is quite simple. The apparent complexity of his behavior over time is largely a reflection of the complexity of the environment in which he finds himself.”
(Simon, 1972, p. 25)

If this assumption was used as the basis for anticipating failures, then the focus would be on the variability of the environment or circumstances and not on the possibility of a failure of the “human component”. Or rather, the possibility of failure would be an attribute of the context and not of the human. More recently, a similar notion has been expressed specifically addressing the issue of error management:

“The evidence from a large number of accident inquiries indicates that bad events are more often the result of error-prone situations and error-prone activities, than they are of error-prone people.”
(Reason, 1997, p. 104)

Interestingly enough, a number of HRA methods can be seen as supporting this view. The classical principle of time-reliability correlation (TRC, cf. Hall et al., 1982) is an expression of the idea that the likelihood of failing in performing an activity is a function of time – although in this case it is time after the onset of an accident rather than time available. A more sophisticated version of the same principle is found in the notion of “error forcing conditions”, although a determining factor here is time available rather than elapsed time (Cooper et al., 1996). The sophistication is due both to the set of conditions that may “force” an error, and the more detailed description of possible error modes. The common feature is that the possibility of performance failure is an attribute of the conditions rather than of the humans.

A closer inspection of a commonly used HRA method such as HEART (Williams, 1988) also reveals the dominance of the circumstances over the individual. Firstly, HEART only refers to the possible failure of an action, but not to specific failure types. Secondly, the characterisation is related to different tasks, which actually means different task conditions. This can be substantiated by a gentle reinterpretation of the basic HEART table, as shown in Table 1.

Table 1: Description of failure types and causes in HEART

Generic tasks	Context or set of circumstances
A. Totally unfamiliar, performed at speed with no idea of likely consequence.	High time pressure, unfamiliar situation
B. Shift or restore system to a new or original state on a single attempt without supervision or procedures	Lack of supervision and procedures
C. Complex tasks requiring high level of comprehension and skill	High task complexity
D. Fairly simple task performed rapidly or given scant attention.	Simple tasks of limited significance
E. Routine, highly-practised, rapid task involving relatively low level of skill.	Routine or highly familiar tasks
F. Restore or shift system to original or new state following procedures, with some checking .	Following a procedure
G. Completely familiar, well-designed, highly practised routine task, oft-repeated and performed by well-motivated, highly trained individual with time to correct failures but without significant job aids.	High-routine task with no time pressure
H. Respond correctly to system event when there is an augmented or automated supervisory system providing accurate interpretation of system state.	Task with monitoring and highly supportive MMI
M. Miscellaneous tasks for which no description can be found.	No specific characteristics

Even allowing for the limited objectivity of the reinterpretation, it is a demonstrable fact that the major source of variability, which determines the likelihood of a failure, is ascribed to the context or circumstances. In other words, the specific working conditions are the signal while the individual human error probability is the noise. The possibility of performance failure is thus an attribute of the conditions rather than of the humans.

3.3 Look To The Performance Conditions

The consequence of this line of argument is that the variability of human performance constitutes the noise rather than the signal. Conversely, the main determinant of performance quality – and specifically of performance failure – comes from the context or the circumstances. The possibility of failure is therefore an attribute of the joint system rather than of any of its components. It follows from this that the anticipation of system failures should concentrate on developing effective ways of describing how joint system performance depends on the conditions rather than on the potential for human failures.

Specifically, predictions should be about how the joint system can lose control of the situation, rather than about whether the human will make a single failure. This would also acknowledge the fact that a human failure is just a single event that requires other conditions to result in an accident – cf. the extended accident model. A practical implementation of this principle can be found in the basic method for performance prediction that is part of CREAM (Hollnagel, 1998a). Here an assessment of the common performance conditions leads to an overall prediction of how likely the operator, hence the joint system, is to lose control. This prediction is made without considering the failure probability for specific actions, or even describing the tasks at the level of component actions.

The CREAM approach assumes that the likelihood of a failure (the proverbial “human error”) depends on the working conditions rather than on the propensity of humans to screw up. It also assumes that it are the system functions and (latent) conditions that determine whether an action failure turns into an accident. In CREAM the basic method for performance prediction starts by characterising the context in terms of a small number of Common Performance Conditions (CPC). Unlike the traditional HRA approaches, the CPCs are assumed to depend on each other in a manner shown by the model shown in Figure 5. In the basic application of the method, the effect of the CPCs alone is sufficient to provide an overall characterisation of the situation, hence of the likelihood of losing control. In the more detailed use, specific error modes – and their probabilities – may be determined, but only on the background of the expected performance conditions. A detailed account of this approach is provided in Hollnagel (1998a).

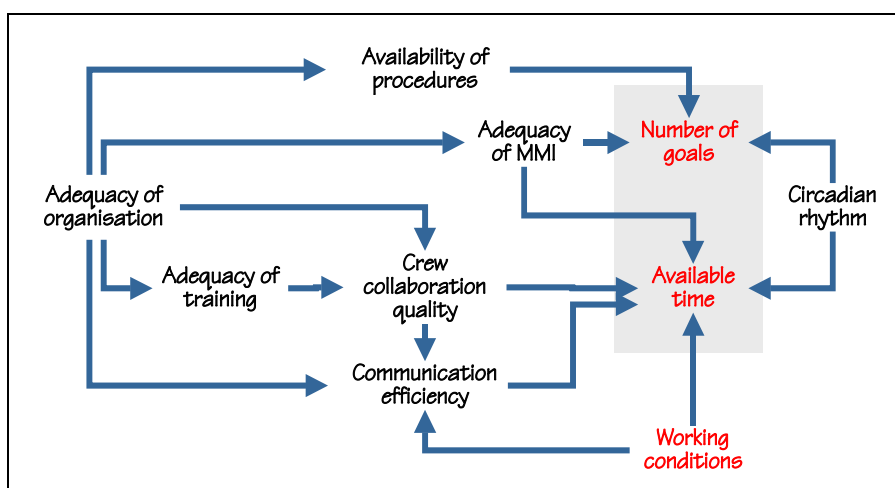


Figure 5: The CPC dependency model in CREAM

4. CONCLUSIONS

I have in this paper tried to argue that there is a need to understand the complexity of joint system performance, and the dynamics of human-machine interaction. Instead of focusing on discrete actions taking

place at single points in time, one should focus on how actions develop over time – on how an event unfolds and how the joint system strives to maintain an equilibrium. It is during this dynamic process that prior events have future consequences, depending on how the conditions change.

A concrete criticism against the established practice is that the repertoire of methods for performance prediction and anticipating failures do not fully reflect the lessons from accident analysis. Despite an impressive and growing amount of evidence, performance prediction remains focused on the notion of “human error”. As argued elsewhere (Hollnagel, 1998a), the concept of “human error” is an artefact of the models and methods that have been used, leads to a view of performance failure as an attribute of the individual – and of human cognition – rather than as an attribute of the context. In order to overcome this bias, we need to develop better models both of how performance conditions affect the likelihood of failure or losing control and of how coincidences can occur and barriers fail. This might also relieve us from hunting after the elusive human error probability and the impossible task of controlling the conditions of observation, and instead look to what is really important – the natural contexts in which people have to work.

5. REFERENCES

- Cojazzi, G. & Pinola, L. (1994). *Root cause analysis methodologies: Trends and needs*. In G. E. Apostolakis & J. S. Wu (Eds.), *Proceedings of PSAM-II*, San Diego, CA, March 20-25, 1994.
- Cojazzi, G. (1993). *Root cause analysis methodologies. Selection criteria and preliminary evaluation* (ISEI/IE/2442/93). JRC Ispra, Italy: Institute for Systems Engineering and Informatics.
- Cooper, S. E., Ramey-Smith, A. M., Wreathall, J., Parry, G. W., Bley, D. C., Luckas, W. J., Taylor, J. H. & Barriere, M. T. (1996). *A technique for human error analysis (ATHEANA)* (NUREG/CR-6350). Washington, DC: US Nuclear Regulatory Commission.
- Dougherty, E. M. Jr., & Fragola, J. R. (1988). *Human reliability analysis. A systems engineering approach with nuclear power plant applications*. New York: John Wiley & Sons.
- Hall, R. E., Fragola, J. R. & Wreathall, J. (1982). *Post-event human decision errors: Operator action trees/time reliability correlation* (NUREG/CR-3010). Washington, DC.: USNRC.
- Hollnagel, E. (1993a). *Human reliability analysis: Context and control*. London: Academic Press.
- Hollnagel, E. (1993b). The phenotype of erroneous actions. *International Journal of Man-Machine Studies*, 39, 1-32.
- Hollnagel, E. (1998a). *Cognitive reliability and error analysis method – CREAM*. Oxford: Elsevier Science.
- Hollnagel, E. (1998b). Context, cognition, and control. In Y. Waern, (Ed.). *Co-operation in process management - Cognition and information technology*. London: Taylor & Francis
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Kirwan, B. (1994). *A guide to practical human reliability assessment*. London: Taylor & Francis.
- Park, K. S. (1987). *Human reliability. Analysis, prediction, and prevention of human errors*. Amsterdam: Elsevier.
- Perrow, C. (1984). *Normal accidents: Living with High-Risk Technologies*. New York: Basic Books.
- Reason, J. T. (1990). *Human error*. Cambridge, U.K.: Cambridge University Press.

Reason, J. T. (1992). The identification of latent organisational failures in complex systems. In J. A. Wise, V. D. Hopkin & P. Stager (Eds.), *Verification and validation of complex systems: Human factors issues*. Berlin: Springer Verlag.

Reason, J. T. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate.

Senders, J. W. & Moray, N. P. (1991). *Human error. Cause, prediction, and reduction*. Hillsdale, NJ.: Lawrence Erlbaum.

Simon, H. A. (1972). *The sciences of the artificial*. Cambridge, MA.: The M. I. T. Press.

Williams, J. C. (1988). *A data-based method for assessing and reducing human error to improve operational performance*. Proceedings of IEEE 4th Conference on Human factors in Power Plants, Monterey, CA, 6-9 June.

Woods, D. D., Johannesen, L. J., Cook, R. I. & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers and hindsight*. Columbus, Ohio: CSERIAC.

This page has been deliberately left blank



Page intentionnellement blanche

Can Human Performance be Addressed Within the Current Safety Assessment Process?

Maarten Boasson

Naval Command and Control System
Anti-Weapon Systems
Universiteit van Amsterdam
Hollandse Signaalapparaten B.V.
P.O. Box 42
7550 GD Hengelo, The Netherlands

Can human performance be addressed within any safety assessment process?

Content

- Context
- Questions
- Observations
- Conclusions

System boundary

- Human operator is not part of a system
 - we design systems; it is presumptuous to suggest we can design human operators
- The interface through which an operator interacts with a system, is part of that system
 - including rules and constraints for usage

Human performance

- Must be studied in relation to the system
- It has many facets
 - reaction time
 - quality of decision, given available information
 - manipulation of controls
 - alertness
 - bias
 any of these can lead to disaster!
- Can human performance be quantified? Some aspects of human performance can such as response time to a given stimulus, but others escape even formal description.

Safety

- A system is intrinsically safe, if under no circumstance a catastrophe is caused by actions in which the system is involved.
- It seems unlikely that such systems can be built!
 - Relative to the definition of catastrophe
- Safety of systems relies on three aspects:
 - correctness of the design
 - correctness of the implementation, and
 - operation within the design limits
 under the assumption of correct specifications.

Limits of our abilities

- It is impossible to predict all possible circumstances a system can be in.
 - We do not generally control the environment:
 - turbulence
 - hijackers
 - imperial to metric conversion
- It is equally impossible to predict all possible system malfunctions, and the associated system behaviours.
 - At least in software intensive systems.
- It is utterly impossible to foresee all possible human actions.
- Is it possible to define all allowed system states?
- If so, can a system be constrained to always be in one of these states?
- Thus, e.g. can faulty operator action be corrected automatically?
- Currently there is no rigorous way to demonstrate correctness of a design; this is true regardless of the engineering discipline involved (but probably more so for software than for other disciplines).
- Establishing that a design is correct w.r.t. a given specification, is a matter of extensive discussion, walkthroughs, etc. Thus, in essence it is a matter of belief and trust.
- Formal checking of conformance between a (certified) design and its implementation, is beyond our abilities.
- At best, extensive testing suggests there are no major implementation errors. Note that software does not really have an implementation stage: the complete design is in itself the implementation. That, unfortunately, does not make it any easier to demonstrate correctness.
- Operation within the design limits requires absence of malfunctions in all of the parts, as well as correct behaviour of the system operator(s).
- Overload in software systems typically occurs as a result of incorrect functioning of either sensing devices (producing more measurements than anticipated), or operators (issuing illegal commands, e.g.).
- Any process aiming at establishing safety of a system, must necessarily contain a large component that relies on human insight, rather than on formal techniques.
- The resulting qualification can therefore not be construed as a guarantee for safety; at best, it provides some measure of confidence that the system is unlikely to fail under circumstances for which it was designed.
- It is questionable whether probabilities given for catastrophic failures of software intensive systems have any useful interpretation.
 - What does the aircraft industry's 10^{-9} mean?
 - The traditional reliability model is unsuitable for software.
- Operator interfaces are part of the design, but operators are not.
- The best that can be done is to specify acceptable operator behaviour under as a wide a variety of circumstances as possible.
 - But we do not know all possible circumstances.
- Providing “natural” interfaces helps
- Operator interfaces can potentially be designed to limit the operator to perform acceptable actions only. This severely reduces the effectiveness of the human operator when the system no longer meets the design constraints. In fact, it reduces the operator to an agent that could have been automated.
 - E.g. an aircraft could be made to refuse execution of an excessively steep dive

- A system can be designed to correct human actions that are considered erroneous (or unsafe), i.e. leading to the system going out of its allowed boundaries. This is similar to refusing illegal commands, but may allow a little more freedom, at greater risk.
- How often can a system recognize such actions? Until we have a formalism for their characterization, it seems difficult for a system designer to make a system sensitive to them.
- Operational procedures and operator training can go a long way in making human behaviour predictable.
- In the limit case, processes developed for assessing system safety, can also be used for human performance w.r.t. safety issues. But then, the operator has been reduced to a finite automaton, and could (should?) have been replaced.
- Operator behaviour not totally governed by operational procedures can hardly be analysed for potential effect on system safety.
- How do we quantify human behaviour?
- The possibility that a human operator will use the interface in unforeseen and dangerous ways, must be taken into account; but how?
 - Note that dangerous situations may well be the result of long chains of interactions between operator and system.
- There is a fundamental conflict between predictability of human behaviour and the ability of man to act intelligently.
- There is a fundamental conflict between predictability of human behaviour and the ability of man to act intelligently.
- Intelligent actions are generally necessary to compensate for system errors (whether due to design faults or material failures).

Conclusion

- For a system to be operated safely, an intelligent human operator is necessary.
- However, a human operator is an intrinsically unsafe component of the <human, system> pair.
- Training and selection are our best friends for improving human performance.
- Quantitative measures for system safety are highly suspect when software is involved.
- Safety assessment can at best give qualitative indications of the likelihood that operators will violate system safety rules.

This page has been deliberately left blank



Page intentionnellement blanche

THEA – A Technique for Human Error Assessment Early in Design

Steven Pocock, Peter Wright, Michael Harrison*

Department of Computer Science

University of York

Heslington, York. YO10 5DD, UK

SUMMARY

Human activity constitutes a major source of vulnerability to the integrity of interactive systems. Wherever human actions are either inappropriate, incorrect, or erroneous, there will be implications for design. This is especially true in high risk endeavours such as commercial air and marine transportation, power production, medical care and space flight. The aim should therefore always be to design an interactive system as resilient to human erroneous actions as possible, and to achieve this as early as possible in the design phase. We present in this paper a formative error assessment technique contributing to the achievement of this goal, known as the Technique for Human Error Assessment (THEA). The method has been applied to several real-world case studies and has demonstrated its suitability in evaluating a design for its vulnerability to human interaction failures which may become problematic once the design becomes operational.

Keywords

THEA, scenario, cognitive failure, error analysis

INTRODUCTION

It has been estimated [2] that approximately 60-90% of all system failures are the direct consequence of human erroneous actions. The concern for safe and reliable performance has understandably been especially high in the nuclear power industry where techniques such as Probabilistic Safety Assessment (PSA) and Human Reliability Analysis (HRA) have been extensively employed. Other methods for assessing the impact of erroneous human actions on interactive systems have since appeared – some qualitative, others quantitative, but it is not intended in this report to review such methods. A brief discussion of some of these can be found in, for example, [7] [8] [2]. The THEA method described in this paper, has its roots in the class of methods of HRA and is designed to inform human-machine interface (HMI) design at an early stage of development.

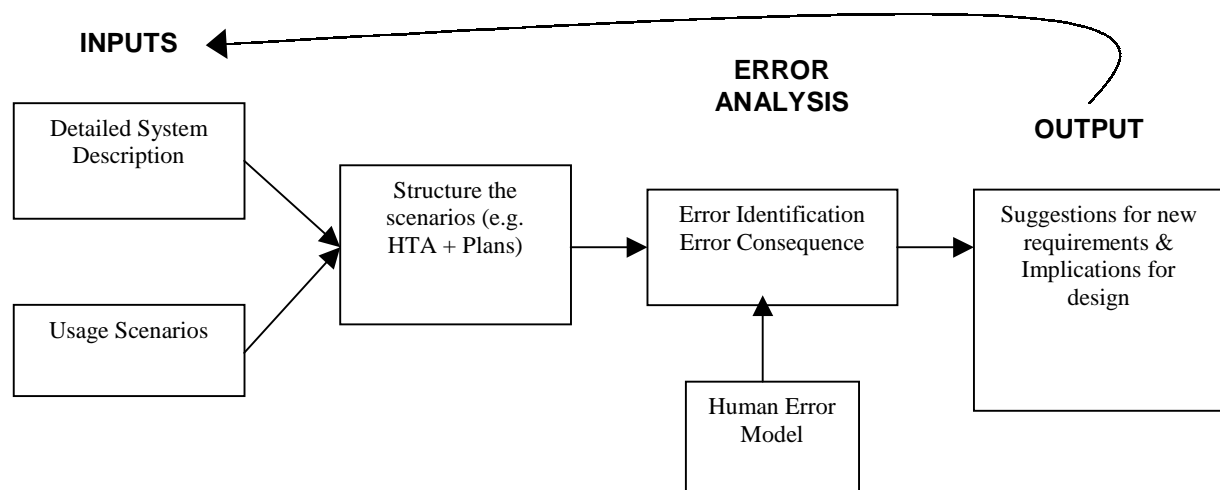


Figure 1: The THEA process

*Current address: DERA, Centre for Human Sciences, Farnborough, Hants GU14 0LX, UK

THEA possesses some similarities with formative evaluation techniques such as Cognitive Walkthrough [9]. In contrast however, THEA aims to consider not only problems with information presentation and feedback but also problems with the planning and execution of actions. THEA also takes a hierarchical view of goals and actions in addition to the sequential perspective of Cognitive Walkthrough. THEA is a strongly suggestive method, guiding the analyst in a structured way to consider areas of a design for potential interaction difficulties. Other methods, such as the *human error identification in systems tool* (HEIST) described in [5], possess similar goals to THEA, except that THEA achieves them with considerably less exertion – eighteen error analysis questions as opposed to 113, which is perhaps why the latter approach has remained largely theoretical. It would certainly be impractical to use without tool support, whereas THEA has the capability of conducting sizeable analyses by means of a prototype tool called ProtoTHEA.

The basic philosophy of THEA views errors as contextualised phenomena influenced by, for example, performance shaping factors. Thus for any method to effectively assess a design for vulnerability to error, it must take account of context. THEA explicitly takes contextual and cultural issues into consideration by means of usage scenarios. In this way it is hoped to elicit the way work is *actually* practiced and not simply how designers *envisage* it as being practiced.

We commence with an overview of THEA, followed by a case study to illustrate the technique.

THEA

The main aim of THEA is to use systematic methods of asking questions and exploring interactive system designs based on how a device functions in a scenario. The purpose of doing this is to provide a systematic and structured means of critiquing a design and developing further requirements [1]. In this way, it is hoped to assist system designers anticipate human interaction failures which may become problematic once a design becomes operational. The technique is intended primarily for use early in the development lifecycle whilst functionality is emerging, and begins with a formal description of the work under analysis. This is achieved by combining two primary inputs consisting of a detailed description of the design under consideration – preferably with domain expert input – and a number of usage scenarios. These inputs, together with the remainder of the THEA process, are illustrated in Figure 1.

Scenarios

THEA views performance failure as an attribute of “cognition in the world” [4], that is to say, of the context or the circumstances which play a fundamental role in its methodology. Applying a communications analogy (op.cit.), performance conditions – or context – may be thought of as the ‘signal’, with erroneous human actions as ‘noise’ superimposed on it. Too little signal and the communication becomes unintelligible. Thus by analogy, with insufficient context, performance failure becomes less meaningful. THEA analyses attempt, through use of detailed scenarios, to capture those complex conditions which result in the human behaving in an unanticipated and unintended manner.

Scenarios should thus comprise not only *actions* which take place in a given situation, but also *contextual factors* which surround the action, allow it to happen, and provide opportunities for “error”. To represent the context as comprehensively as possible, a scenario template in [1] incorporates the following information:

1. Agents
 - The human agents involved and their organisation
 - The roles played by the humans, plus their goals and responsibilities
2. Rationale
 - Why is the scenario interesting?
3. Situation and Environment
 - The physical situation in which the scenario takes place
 - External and environmental triggers, problems and events that occur in this scenario

4. Task Context
 - What tasks are carried out?
 - What formal procedures exist, and are they followed as prescribed?
5. System Context
 - What devices and technology are involved? What usability problems might they possess?
 - What effect can users have?
6. Action
 - How are the tasks carried out in context?
 - How do the activities overlap?
 - Which goals do actions correspond to?
7. Exceptional circumstances
 - How might the scenario evolve differently?
8. Assumptions
 - What, if any, assumptions have been made?

Principal sources for scenario elicitation include:

- Experience with earlier versions of the system. ‘Top-down’ designs are relatively infrequent and previous versions usually have associated reports highlighting problem areas;
- Incident and accident reports;
- Frequent conditions and normal operation;
- Where technology changes. This is the principal source for the case study presented in this paper;
- Where concepts change. For example, changing from conventional air traffic control to Datalink.

Finally, we want to know how many scenarios will be required to capture the usage context in sufficient detail. The answer is really reliant upon expert judgement as to when a ‘good enough’ coverage has been achieved, and for this reason it is highly desirable to have at least one domain expert involved in the scenario construction process.

Goal Decomposition

To structure and interpret information contained in scenarios, Hierarchical Task Analysis (HTA) is a practical – but by no means the only – way of achieving goal decomposition. It is hierarchical because task goals are broken down into a structure of sub-goals which must first be achieved before the top level goal can be satisfied. In this way we can describe operators’ tasks in terms of the goals and sub-goals to be achieved and the actions used to achieve these goals. Plans are appended to each task to describe the flow of control through the task and detailing *how* the sub-goals and actions within a task are combined to satisfy the higher level goal.

Task descriptions, while good at describing what a user has to do and know, is less adept at describing how an interface might respond to a user’s inputs. THEA presumes that some notion of causality can be used to explore the interaction between for example, a display and other perceptual cues, operator memory requirements, and other aspects of the design. A set of behavioural analysis guidewords (omission, commission, and so on) is employed, based on a control model of operator-system interaction [6]. These can trigger questions about the extent to which, for example, a display is able to support goals and plans, or to consider how apparent it would be for an operator to perform an appropriate action. We believe this affords a means of linking task and system descriptions more directly, and forms the basis of the THEA error analysis phase.

Error Analysis

The foregoing steps identify a number of factors facilitating an understanding of the context in which human actions – and therefore erroneous actions – take place. We are now in a position to draw these strands together in the analysis phase which helps identify where HMI error may be problematic.

The analysis adopts a structured questionnaire-, or checklist-, style approach, referred to in [1] as the “Cognitive Error Analysis”. This is based on failures (Table 1) that are possible in Norman’s execution-evaluation cycle model of human information processing [6].

Table 1: Examples of cognitive failure

Stage	Cognitive failure
Goals	Lost/Unachievable/Conflicting No triggering/activation Triggering/activation at wrong time, or wrong goal activated
Plans	Faulty/Wrong/Impossible
Actions	Slip/Lapse
Perception/ Interpretation	Failure to perceive correctly Misinterpretation

The error analysis poses questions about the scenario to reveal areas of design where cognitive failures may occur, and assess their possible impact on the task or system being controlled. A simple example might be the high level goal of photocopying a sheet of paper. One of the THEA analysis questions asks whether the goal can be accomplished without all its sub-goals being correctly achieved. The analyst would typically answer (in the case of most photocopiers) “yes” since it is entirely possible to walk away with your copy but leave the original document and/or copier card in the machine. The sub-goal has thus been lost and a ‘post-completion’ error has occurred. A full list of the THEA error analysis questions can be found in Appendix A.

There will be occasions when no obvious behavioural manifestations are evident. For example, if an operator is presented with conflicting goals, this may itself be a ‘manifestation’ of the problem which, if serious enough, may require a design solution to be found.

Exactly how the analysis is carried out is largely a matter of choice, but the two envisaged methods are:

1. Follow the goal hierarchical structure from top to bottom asking each question about each goal or action;
2. Select parts of the scenario where potential problems are anticipated, then conduct a detailed analysis of behavioural error and impact where appropriate.

Clearly the first option is the most thorough and is recommended for new designs. Understandably it is probably going to be lengthy and time consuming but also likely to uncover a greater number and range of concerns.

Recording the results

Whichever approach is adopted, the analysis results may be recorded according to project requirements. We have found, however, that a tabular format provides a practical way of presenting the information. **Table 2** shows a typical arrangement, while table 4 provides an example:

Table 2: Tabular format for recording EA results

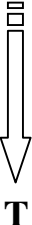
Question	Causal Issues	Consequences	Design Issues
Question identifier as an aid to traceability	Issues raised by analyst	Consequences of the causal issue	Notes, suggestions, comments, re-design ideas

From our own case study work, sometimes involving large and complicated scenarios, we identified a need for tool support to assist the analyst with the entry, handling, and storage of information associated with a project. This resulted in the development of ProtoTHEA, a prototype tool where the scenario, HTA, and error analysis details of a particular project can be entered via a graphical user interface. All information is held in a database and an output, in the form of ‘failure state profile’ charts (adapted from [8]), is automatically obtained for each scenario. Appendix B shows a typical ProtoTHEA HTA and error analysis example, as well as a failure state profile chart.

APPLICATION OF THEA – CASE STUDY

We now illustrate a practical application of THEA by means of a case study, based on information collected from flight crew, involving a change of technology on the flight deck of a fisheries reconnaissance aircraft. A major change between the old and the new flight decks concerns the crew complement being reduced from three people to two, the flight engineer being replaced by computerised technology. The scenario involves a situation where the activities of the flight engineer would, on the old flight deck, be particularly significant. We deal with emergency conditions rather than normal operation, but since the tasks in themselves are fairly straightforward and do not involve much decision making, the crew activities involve more knowledge intensive activities such as fault diagnosis.

Table 3: Scenario timeline showing actions – some conflicting – performed by each agent



System status	Pilot flying (PF)	Pilot not flying (PNF)	Information sources	System response
Engine 3 fire warning	Throttle 2 max. Press master warning Throttle 1 idle	Close bomb bay doors Flaps 0 Rudder trim Warn crew	Airmanship Airmanship	Select ENG ECAM page
Engine 4 fail warning	Throttle 1 max	Throttle 3 close LP cock 3 shut Fire ext 3; shot 1	Engine 3 fire drill	Start engine
	Navigate safe exit route			

Situation and environment

The starting condition involves a four-engine fisheries patrol aircraft at low level over water, photographing a fishing vessel. To conserve fuel, the aircraft is flying on engines 2,3,4 only. Engine 1 (leftmost) has been closed down for fuel economy reasons. The aircraft suffers a massive bird strike on the right side. As a result of bird ingestion to engines 3 and 4, both engines fail producing engine failure and engine fire warnings. The engine problems will cause the failure of the generators in these engines, which will in turn lead to the remaining generators being overloaded, resulting in a series of warning or cautions being signalled after a short delay.

Actions in context

As we discussed earlier, one of the principal components of a scenario is a description of the actions which take place. An HTA may be employed, but it is not always necessary. If, for example, interaction with the system of interest is relatively simple, then it is probably sufficient to simply identify the goals users have, and write down a list of the actions necessary to achieve the goals. If the interaction is more complex, then a more formal approach for capturing tasks and goals, such as HTA, may be needed. For this scenario, we adopt the former approach since it is not the intention here to produce a fully worked example, rather to give a flavour of how the technique may be used.

In Table 3 we show some of the crew and ‘system’ actions in the early stage of the scenario, with time flowing downwards. What is interesting is that one can observe both pilots conducting possibly contradictory actions at the same time – the PF is attempting to restart engine 1 to produce more thrust, while the PNF is shutting down the faulty engines i.e. *reducing* thrust. However, what this diagram does *not* show are links between actions and the surrounding context, which is a main reason for thinking about scenarios in the first place. To accommodate this, Table 3 may be modified to include the goals – derived from the task analysis – to which they are directed. Figure 2 shows a goal structured action sequence for our scenario with time now represented qualitatively along the horizontal axis. The same actions as before are shown but, in addition, the goals that drive the interaction – as well as triggers that bring the goals into being – can be seen. Presenting scenario actions in this way illustrates a number of features not immediately evident in, for example, a traditional HTA. In particular, Figure 2 shows which goals and tasks become active, and active concurrently in the scenario, as well as which actions are related by being directed towards the same goals. These are not present in the simple event listing of Table 3 which makes no mention of goals.

Analysis example

An illustration of how the analysis is conducted is shown in Table 4. We have selected only two of the questions from the full cognitive error analysis question list (see Appendix A) which are particularly pertinent to this scenario, namely:

G1 – The mechanisms which trigger or activate goals, and

G3 – The potential for conflicting goals.

Asking question G1 yields a number of possible answers since different collections of goals have different triggering properties. Some are fairly innocuous and do not suggest potential problems (e.g. “Shut down engine” is triggered quite directly by a warning), whereas others are less directly triggered and may be more prone to being omitted (e.g. “Engine 3 cleanup”). A full version of the analysis is provided in [1].

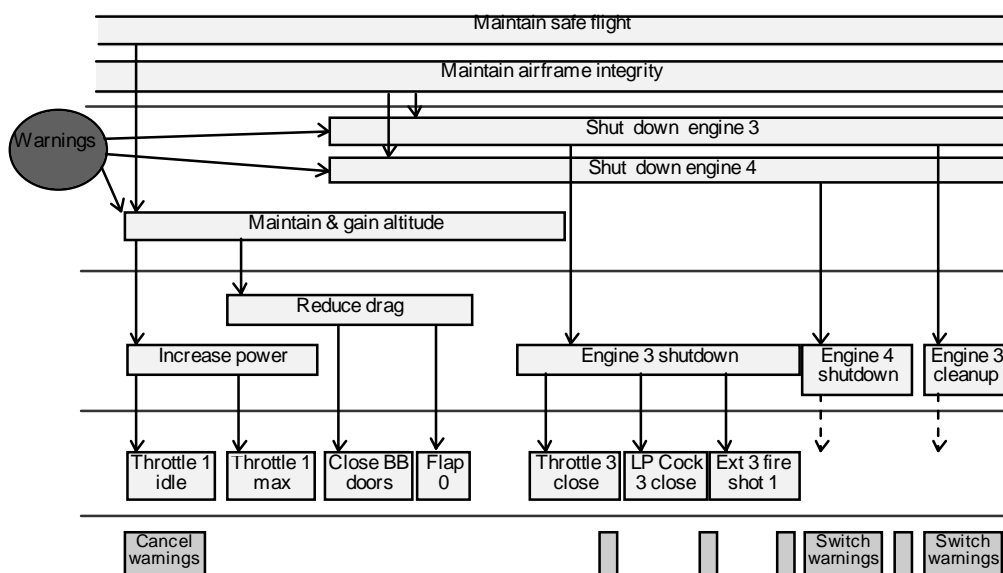


Figure 2: Hierarchical goal structuring of scenario actions

Table 4: Example application of error questionnaire

Question	Causal Issues	Consequences
G1 (Triggers, task initiation)	<p>Many goals triggered fairly directly (e.g. “Shut down engine 3”)</p> <p>Timing of lower level goals arises as a combination of triggering and group decision making (e.g. Engine 3 shutdown)</p> <p>Some goals rely on general Airmanship skills for their activation (e.g. power, drag)</p> <p>Some goals are poorly triggered, especially if there are several goals with only a single trigger on the display (e.g. “Engine 4 shutdown” or “Engine 3 cleanup”).</p>	<p>Main behavioural consequence is that triggers for cleanup actions exist in the display, but are removed when other tasks intervene – switching to “Engine 4 shutdown” removes indications for “Engine 3 cleanup”).</p> <p>It is also possible that “Engine 4 shutdown” or “Engine 3 cleanup” might be omitted or delayed.</p>
G3 (Goal conflicts)	<p>Goals to increase power and Engine 3 shutdown are in conflict (although this is inevitable)</p>	<p>Resolving the conflict satisfactorily requires negotiation between PF & PNF. The time required for this may lead to a non-optimal (too late) decision.</p>

When performing a full analysis, causal issues raised producing noteworthy or problematic consequences are documented in the ‘consequences’ column. Entries for certain questions might be left blank, indicating that the question did not appear to reveal any interesting insights. A third column could also be added entitled “Design Suggestions”. Thus we might add to G3 “Attempt to design out conflicts or give participants the resources to resolve them”, and so on.

Finally, it is worth mentioning that certain ‘keywords’ (omission, commission, etc.) will not make sense in the context of every scenario. For example, a ‘repetition’ error is not-applicable to an aircraft’s take-off sequence. In other cases, physical constraints may make it impossible, or it would be hard to imagine how such deviations might occur.

THEA & QUANTIFICATION

The primary output of THEA is a description of a number of problem areas associated with a design and its operation which may be the cause of interaction errors. These are intended to assist designers reason about errors at the early stages of a design before it becomes impractical or prohibitively expensive to effect a longer term design change or implement shorter term procedural ‘fixes’ or limitations.

Unlike some hazard identification methods such as hazard and operability studies (HAZOP), THEA does not directly identify hazards per se but instead addresses the causal factors which contribute to them. That is, it does not provide quantitative estimates of the likelihood of human erroneous actions. This is not to say that the method proscribes the use of supplemental quantification where useful or necessary. For example, THEA has been supplemented in certain case studies by the Human Error Assessment and Reduction Technique (HEART) [10]. This is a task-based approach utilising a database of error probabilities, and relying on the application of a simple algebraic formula to a chosen generic task and weighted error producing conditions (EPCs). It has demonstrated its usefulness in supporting THEAs qualitative output by allowing us, where a number leads to a concern, to ask:

- Have we chosen the wrong generic task?
- Have we chosen inappropriate EPCs?
- Have we weighted the EPCs disproportionately?

In this way, our assumptions, both qualitative and quantitative, may be reflected upon and revised if necessary. The advantage of a supplemental method such as HEART is that it is readily understandable by all interested parties and is a way of supporting dialogue about human reliability estimates.

While numbers may be useful, it is important to be clear how they are intended to be used. We must also be quite clear what they represent and to whom. For example, the ‘traditional’ engineering view regards numbers as representing real values of probabilities which may be combined and manipulated arithmetically. In our experience, numbers represent broad categories of risk and serve as ‘tokens’ for the negotiation of concerns (“Do we have a problem?” or “I think your estimate for this error is unrealistic”). That is to say, numbers should not be treated as objective truths but rather as starting points for discussion. Superficially, qualitative and quantitative predictions are different outcomes, but it will be appreciated that they are actually opposite sides of the same coin. As Hollnagel [3] points out:

“Quantification can only be done for something that has been clearly identified and described, and this description must necessarily be qualitative. Quantities must be quantities of something, and that something must be previously described.” (p.80)

Whilst it may be argued that a quantitative approach is necessary to support and satisfy conditions of, for example, a Probabilistic Safety Assessment (PSA) or a specific customer requirement (“No single failure shall have a catastrophic or critical hazardous consequence in every 10^9 hours”), it is uncertain whether, or to what extent, such an approach actually matches reality. All quantitative methods are ultimately based on a qualitative description and some underlying model. It follows that if any of the descriptive steps are lacking, the outcome of any numerical analysis will necessarily be incomplete no matter how refined the quantification process.

DISCUSSION

This paper has described a formative error analysis technique, THEA, for analysing system vulnerability to erroneous human actions. One of the most important antecedents of the THEA error analysis process is gaining an understanding of how the system being examined will be used in practice. We formulate ‘usage scenarios’ to furnish us with context of use – the circumstances or conditions under which an event occurs – to elicit how work will *actually* be performed as opposed to how it is *envisaged* it will be performed.

It is highly desirable to carry out an analysis early in the design process before adverse consequences are encountered at ‘the sharp end’. THEA anticipates, through design critique, interaction failures which may become problematic once a design is operational. In such a way it can assist in developing further requirements before a design becomes ‘rigid’ and excessively difficult or expensive to modify. We differentiate between cause and consequence since incorrect operator actions and assessments are treated as the *starting* point for analysis rather than the conclusion – they are recognised as *symptoms* rather than causes. In this predictive role, causes are the initiating events and manifestations are the possible outcomes. Of course, THEA works equally well for retrospective analyses of extant designs. A recent case study employed the technique to appraise a system where specific erroneous operator actions would result in serious consequences. THEA highlighted system design issues contributing to such performance as well as providing an assessment of possible consequences. Our results supported the clients’ numerical analysis thus affording a more confident design assessment. In addition, the case study facilitated convergence of practitioners and human factors personnel through the exchange of ideas and techniques. This helped overcome what Hollnagel refers to in [3] as “the *conceptual impuissance or abstruseness*”.

We have found from experience that, although no special expertise is required to carry out the error analysis procedure, input to the process by domain experts significantly expedites its completion. Additionally, tool support offered by ProtoTHEA has demonstrated an ability to manage large and complex case studies. Whether the ‘traditional’ or tool-assisted approach is employed, the emphasis of THEA is on functionality and practicality, both ably demonstrated in recent case study work.

REFERENCES

1. Fields, B., Harrison, M., & Wright, P. (1997) THEA: Human Error Analysis for Requirements Definition. YCS 294 (1997)
2. Hollnagel, E. (1993) *Human Reliability Analysis: Context and Control*. Academic Press Limited.
3. Hollnagel, E. (1998) *Cognitive Reliability and Error Analysis Method CREAM*. Elsevier Science Ltd.
4. Hollnagel, E. *The human factor in system reliability: Is human performance predictable?*, in *Anticipating failure: What should we make predictions about?* 1999, NATO RTO HFM Workshop: Certosa di Pontignano, Siena, Italy.
5. Kirwan, B. (1994) *A guide to practical human reliability assessment*. Taylor and Francis.
6. Norman, D.A. (1988) *The psychology of everyday things*. Basic Books.
7. Reason, J. (1990) *Human error*. Cambridge.
8. Reason, J. (1997) *Managing the Risks of Organizational Accidents*. Ashgate Publishing Ltd.
9. Wharton, C., et al. *The Cognitive Walkthrough method: A practitioner's guide*, in *Usability Inspection Methods*, R.L. Mack and J. Nielsen, Editors. 1994, John Wiley and Sons, Inc. p. 105-140.
10. Williams, J. (1986) HEART - A proposed method for assessing and reducing human error. *Proceedings of 9th Advances in Reliability Technology Symposium*. (1986, University of Bradford).

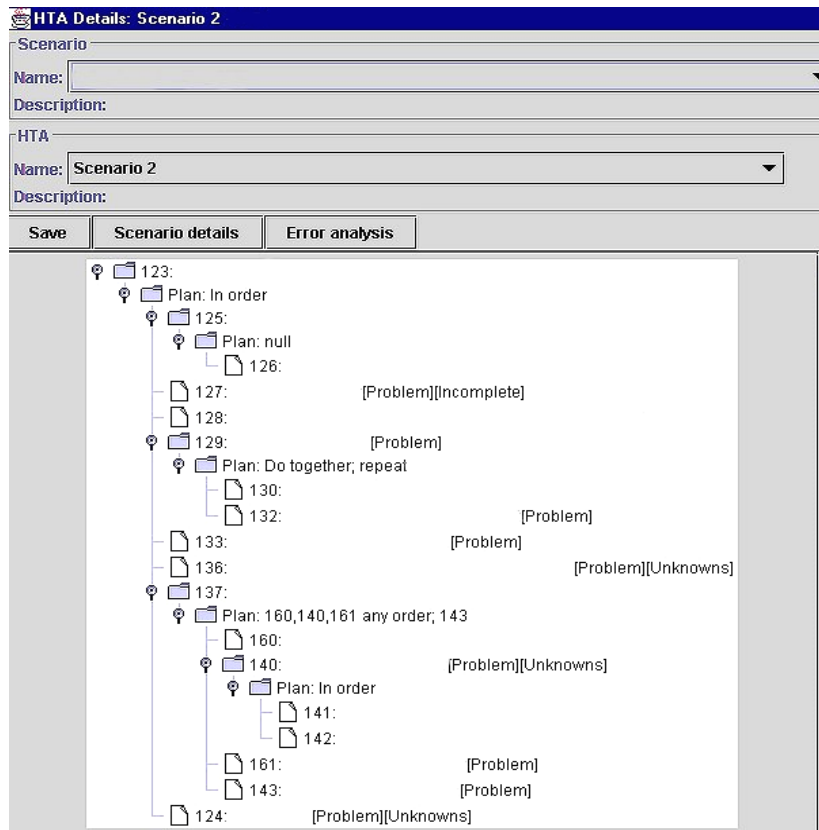
APPENDIX A – THEA error analysis questions

Questions	Consequences	Examples & design questions
Goals, Triggering and initiation		
G1. Are items triggered by stimuli in the interface, environment, or task?	If not, goals (and the tasks that achieve them) may be lost, forgotten, or not activated, resulting in omission errors.	Are triggers clear and meaningful? Does the user need to remember all the goals?
G2. Does the user interface “evoke” or “suggest” goals?	If not, goals may not be activated, resulting in omission errors. If the interface does “suggest” goals, they may not always be the right ones, resulting in the wrong goal being addressed	E.g.: graphical display of flight plan shows pre-determined goals as well as current progress.
G3. Do goals come into conflict?	If so additional cognitive work (and possibly errors) may result from resolving the conflict. If the conflict is unresolvable, one or more goals may be lost, abandoned, or only partially completed.	Can attempt to design out conflicts or give participants the resources to resolve them.
G4. Can a goal be achieved without all its “sub-goals” being correctly achieved?	The sub-goals may be lost (resulting in omissions).	E.g.: goal of photocopying achievable without sub-goal of retrieving card.
Plans		
P1. Are there well practised and pre-determined plans?	If a plan isn’t well known or practiced then it may be prone to being forgotten or remembered incorrectly. If plans aren’t pre-determined, and must be constructed by the user, then their success depends heavily on the user possessing enough knowledge about their goals and the interface to construct a plan. If pre-determined plans to exist and are familiar, then they might be followed inappropriately, not taking account of the peculiarities of the current context.	
P2. Can actions be selected in-situ, or is pre-planning required?	If the correct action can only be taken by planning in advance, then the cognitive work may be harder. However, when possible, planning ahead often leads to less error-prone behaviour and fewer blind alleys.	
P3. Are there plans or actions that are similar to one another? Are some used more often than others?	A more common but similar plan may be confused for the intended one, resulting in the substitution of an entire task or sub-task.	
Performing actions		
A1. Is there physical or mental difficulty in executing the actions?	Difficult, complex, or fiddly actions are prone to being carried out incorrectly.	
A2. Are some actions made unavailable at certain times?		
A3. Is the correct action dependent on the current mode?	Creates a demand on the user to know what the current mode is, and how actions’ effects differ between modes. Problems with this knowledge can manifest themselves as a substitution of one logical action for another.	
A4. Are additional actions required to make the right controls and information available at the right time?	The additional goals may be lost (resulting in omissions) and users will be unable to carry out the main goals. The overall effect may be to cause confusion and disorientation for the user.	

Perception, Interpretation and evaluation		
I1. Are changes (resulting either from user action or autonomous system behaviour) perceivable?	If changes are not perceivable, the user must retain a mental model of the system state. Particularly problematic if changes happen autonomously.	
I2. Are the effects of actions perceivable immediately?	If there's no feedback that an action has been taken, the user may repeat actions.	
I3. Does the item involve monitoring, vigilance, or continuous attention?	The user's attention can easily be diverted away from monitoring tasks, meaning that changes that confirm goals achievement (leading to repetition of actions or carrying out actions too late) or that trigger new goals may be missed (resulting in omission of the associated actions).	
I4. Can the user determine relevant information about the state of the system?	If not, the user will have to remember the information they require, thus making it prone to being lost or recalled incorrectly .	
I5. Is the relation of information to the plans and goals obvious?	If the relationship to plans isn't clear, then a source of feedback about correct execution of the plan, and therefore a factor that mitigates against error, is lost. If the relationship to goals is unclear, then the user may be unaware of when a goal is achieved, leading to termination of a sub-task too early or too late .	
I6. Is complex reasoning, calculation or decision making involved?	If cognitive tasks are complex, they may be prone to being carried out incorrectly , to being the cause of other tasks carried out too late , or to being omitted altogether.	
I7. Is the correct interpretation dependent on the current mode?	Creates a demand on the user to know what the current mode is, and to how the appropriate interpretation of information differs between modes. Problems with this knowledge can manifest themselves as a substitution of one logical information item for another.	

APPENDIX B – ProtoTHEA example: HTA and error analysis extract

The diagrams below show typical extracts from the ProtoTHEA tool. For the HTA in Screenshot 1, specific tasks have not been labelled for clarity, permitting illustration of feedback to user as to the status of each task. This enhances traceability and completeness. The error analysis extract in Screenshot 2 shows a typical screen presented to an analyst, demonstrating the questionnaire nature of the process. All respondent data is stored automatically, and the resultant failure state profile chart for each scenario is shown in Screenshot3.

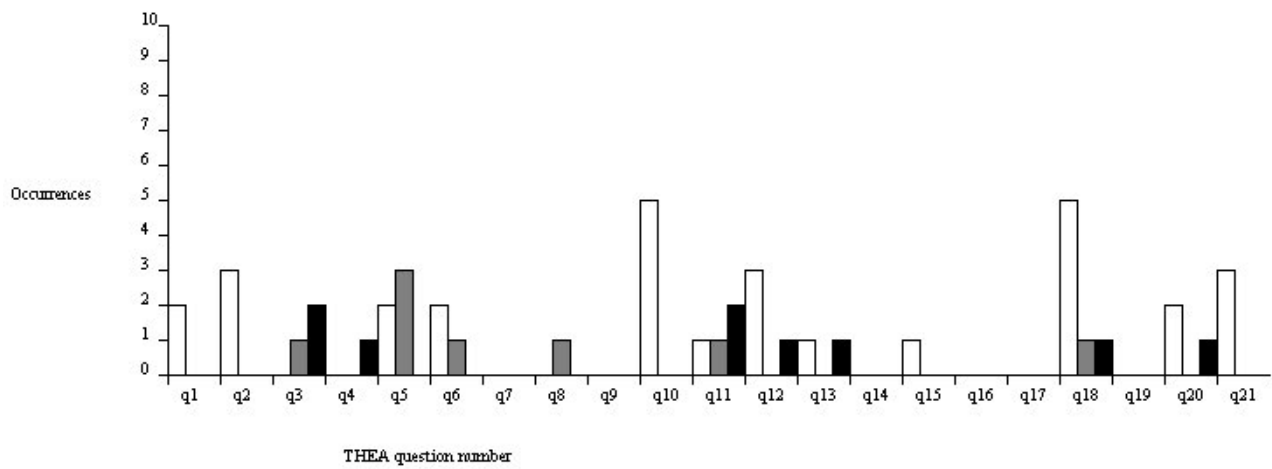


Screenshot 1 – Hierarchical Task Analysis (HTA) extract

Screenshot 2 - Error Analysis extract

APPENDIX B (continued)

HTA Name: Scenario 2



Screenshot 3 – Failure state profile chart for Scenario

This page has been deliberately left blank



Page intentionnellement blanche

Human Reliability in Civil Aircraft Inspection

C. G. Drury

State University of New York at Buffalo
Department of Industrial Engineering
342 Bell Hall
Buffalo, NY 14260
USA

Abstract:

Inspection of structures, systems and engines is an important part of ensuring continued airworthiness of the civil aircraft fleet. This paper describes the airworthiness assurance system and considers applicable bodies of knowledge which help understand and predict aircraft inspection performance. Two examples of recent studies of aircraft inspectors are used to illustrate the extra depth and breadth of understanding available where such knowledge is applied to these tasks. It is concluded that perhaps we have two separate roles: to predict performance and to improve it. Quantitative prediction will never be complete, but better estimates of inspector variability help us set more realistic inspection intervals. However, for improving aircraft inspection tasks we should concentrate on broader contextual factors, despite our inability to quantify some of these effects.

Human Reliability Issues in Aircraft Inspection:

Civil Aviation is growing at over 3.34% per year, and the total annual passengers in many developed countries is comparable to the country's population. Flying is a relatively safe activity, but one whose failures are dramatic and highly publicized. Thus, the exposure to risk is seen by the population as high and increasing. As airline growth increases, the prediction (Boeing, 1997) is for increasing crashes, up to one per week in 2015 unless the current incident rate is decreased.

Of the most visible crashes, known as hull-loss accidents, the fraction with maintenance or inspection as a contributing factor has been about 20% historically (Boeing 1997), but the rate has been increasing in recent years. Thus, inspection and maintenance errors have been seen recently as a major airworthiness emphasis (e.g. Gore Commission Report, 1997).

This paper considers one aspect where human reliability plays a crucial role, that of inspection. The work reported here is the outgrowth of several initiatives by regulatory bodies, primarily the Federal Aviation Administration (FAA) in the USA, Transport Canada and the UK's Civil Aviation Administration (CAA). These range from reliability measurement of inspection tasks to the use of Crew Resource Management (CRM) techniques in maintenance and inspection activities. The aim of this paper is to consider the findings of these initiatives and other applicable human factors knowledge in the domain of aviation maintenance. What can each contribute to improving system reliability? What are lessons for other highly-regulated safety-critical systems which have an inspection component? To do this, we first present an overview of the system for inspecting and maintaining aircraft, and then summarize the findings from contributing fields.

The System:

Airworthiness of civil aircraft depends upon a process by which a team composed of aircraft manufacturers, regulators and one or more airlines predict possible system failures. This process, Maintenance Steering Group 3 or MSG-3, considers possible failure pathways (for example in structures, controls, avionics) and for each pathway determines a recovery strategy. For structural failure, this may be replacement after a fixed service life, regular inspection to assure detection, or an indication to crew of the malfunction. In this paper the concern is with the reliability of the primary failure recovery system for aircraft structural inspection: regular inspection to assure detection.

Failure modes of aircraft structures can be cracks, corrosion, fastener/bonding failure or deformation beyond the plastic limit. Inspection systems are designed to detect all of these in a timely manner, i.e. before the failure has a catastrophic effect as structural integrity. For example, crack growth rates can be predicted probabilistically from material properties and applied stresses, so that the MSG-3 process can schedule inspections before a potential crack becomes dangerous. However, the detection system has certain limits on size crack that can be detected, so that MSG-3 typically schedules several inspections between the time the crack becomes detectable and the time it becomes dangerous. If too many inspections are scheduled, the costs are driven up in a highly-competitive industry, and the risk of collateral damage is increased due to the inspection process itself. Conversely, if too few inspections are scheduled, the probabilistic rate of the crack growth prediction process may combine with the probabilistic nature of the detection process to cause dangerous cracks to remain undetected. Spectacular failures of this inspection process have occurred both for aircraft structures (Aloha incident, Hawaii 1988) and engine components (Pensacola incident, Florida 1997).

The MSG-3 process thus requires quantitative data on inspection reliability to function correctly. In addition, no rule-based prediction system can foresee all possible malfunctions, so that once an aircraft is in service, regular detailed inspections are made of the whole structure to discover any unexpected cracks. When such “new” cracks are found, the information is typically shared between manufacturers, operators and regulators in the form of supplementary inspections. Similar considerations apply to other failure modes such as corrosion.

This whole reliability assurance process thus rests upon an inspection system which checks both points where malfunctions are expected and points where they are not expected, for a variety of malfunctions. For good reasons, human inspectors are part of this inspection system, so that human inspection reliability is an essential element in ensuring structural integrity, and hence airworthiness. The rest of this paper considers bodies of knowledge and data from three sources which should be applicable to human inspection reliability. Parts of this material have been reviewed previously (Drury and Spencer, 1997) to which the reader is referred for further details and references.

The Inspection Task:

The inspection task implied above combines two goals: detection of expected malfunctions and detection of unexpected malfunctions. Neither detection is particularly easy or particularly rapid, so that inspection can be a difficult and time-consuming task. In some ways inspection can be classified as an ill-structured task (Wenner, 1999) because there is no simple step-by-step procedure which will ensure success, and because there is usually no knowledge of task success available during the task. Finding (n) malfunctions in a structure still leaves an unknown number (hopefully zero) potentially undetected.

In addition, inspection is typically scheduled at the beginning of an aircraft's maintenance visit so that malfunctions can be detected early and their repair scheduled to overlap in time with other maintenance activities. As airlines streamline their parts inventory to reduce holding costs, the lead time for replacement components can increase, again pressuring the inspection system to ensure early detection. Aircraft typically arrive following scheduled service, i.e. after the last flight of the day. Following opening up and cleaning processes, maximum inspection resources are committed to the initial inspection. In practice this means inspectors working overtime, even double shifts, starting with a night shift, under some implied pressure for early detection. Human inspection reliability may not be optimal under these conditions.

The inspection task itself is classified in aviation as either Visual Inspection or non-destructive inspection. Regulatory bodies have issued formal descriptions of both of these tasks (e.g. Bobo (1989) for the FAA), and both have somewhat different characteristics in aviation

Non-destructive inspection (NDI) comprises a set of techniques to enhance the ability to detect small and/or hidden malfunctions. One set of NDI techniques are those which enhance what is essentially still a visual inspection task, for example X-ray, fluorescent particle, magnetic particle or D-sight. They show cracks which are very small (fluorescent particle) or hidden within other structures (X-ray). Apart from the steps necessary to ensure a good image, they have many of the human interface characteristics of visual inspection. The other

set of NDI techniques are focused on specific malfunctions in specific locations, e.g. eddy current, ultrasound. For this reason, they are only useful for detection of malfunctions already predicted to exist. In practice, such NDI techniques are much more proceduralized than visual inspection or NDI techniques which contain a human visual inspection component.

Visual inspection is much more common, comprising 80% of all inspection Goranson and Rogers (1983). It consists of using the inspector's eyes, often aided by magnifying lenses and supplementary lighting, as the detection device. Inspectors must visually scan the whole structure of interest, typically using portable mirrors to examine areas not directly visible. Whether the task is categorized as Visual Inspection or NDI, its aim is to detect flaws (indications) before they become hazardous. Next we consider the bodies of knowledge potentially applicable to aircraft inspection reliability. This section is adapted from Drury and Spencer (1997).

Applicable Knowledge 1. NDI Reliability:

Over the past two decades there have been several studies of human reliability in aircraft structural inspection (Rummel, Hardy and Cooper, 1989; Spencer and Schurman, 1995; Murgatroyd, Worrall and Waites, 1994). All of these to date have examined the reliability of Non-Destructive Inspection (NDI) techniques, such as eddy-current or ultrasonic technologies.

From NDI reliability studies have come human/ machine system detection performance data, typically expressed as a Probability of Detection (PoD) curve, e.g. Spencer and Schurman (1995). This curve expresses reliability of the detection process (PoD) as a function of a variable of structural interest, e.g. crack length, providing in effect a psychophysical curve as a function of a single parameter. Sophisticated statistical methods (e.g. Hovey and Berens, 1988) have been developed to derive usable PoD curves from relatively sparse data. Because NDI techniques are designed specifically for a single fault type (e.g. cracks), and much of the variance in PoD can be described by just crack length, the PoD is a realistic reliability measure. It also provides the planning process with exactly the data required, as remaining structural integrity is largely a function of crack length.

Both the FAA (*National Aging Aircraft Research Program Plan*, 1993, p. 26, p. 35) and the Air Transport Association (ATA) have recognized the need for equivalent studies of the reliability of visual inspection as a research priority.

Applicable Knowledge 2. Industrial Inspection:

Human factors analyses of inspection tasks have been published since the 1950's and 1960's with a steady evolution of approaches. Early studies (e.g. Thomas and Seaborne, 1961) tended to be rich and holistic descriptions of inspection tasks. They focused on some of the unique perceptual cues used by experienced inspectors. These showed for example that inspectors organize their perspectives so as to enhance subtle task relevant visual or auditory cues and suppress what a novice would perceive as salient cues. This tradition of description has occasionally resurfaced (Biederman and Shiffar, 1987; Dalton, 1991) but has been largely replaced by more quantitative studies.

The next wave of work measured human performance in a variety of inspection tasks, typically in terms of the two possible errors: missed defects and false alarms. Reviews of this work are readily available (Drury, 1992; Megaw, Alexander and Richardson, 1979). Table 1 classifies some of the factors found to affect inspection performance, using ICAO's SHELL model (ICAO, 1989). Following such studies, and indeed overlapping them, were model-oriented studies treating inspection as either a signal detection task (Harris, 1969; Drury and Addison, 1973) or a visual search task (Kundel; 1975; Drury, 1990). The advantage of such approaches is that they can use the underlying models to predict which variables are most and least likely to affect inspection performance. They also allow succinct descriptions of tasks and task performance, potentially leading to quantitative models. For example, NDI studies of aircraft inspection often provide Relative Operating Characteristic (ROC) curves relating miss rate to false alarm rate for a given defect type.

An inspection model combining search and decision (Drury, 1975) can also be helpful in understanding the inspector's tasks in inspection. This model, summarized in Figure 1, shows an inspector searching an item by repeated fixations of small areas. If an indication (potential defect) is found, a decision task takes place to determine whether the indication should be classed as a reject. If not, or if the fixation found no indications, search continues. The inspection task stops (or moves to the next item) when there is no further time left for inspection, either because of the inspector's stopping policy or external pacing of inspection. This model allows us to specify the variables affecting each stage. Thus, peripheral visual acuity should affect fixation area and thus, search performance (Courtney, 1984). Conversely, the decision stage should be affected by cost and probabilities of the decision outcomes (Chi and Drury, 1998). Overall, this model has been useful in interpreting the speed/ accuracy tradeoff in inspection (Drury, 1994).

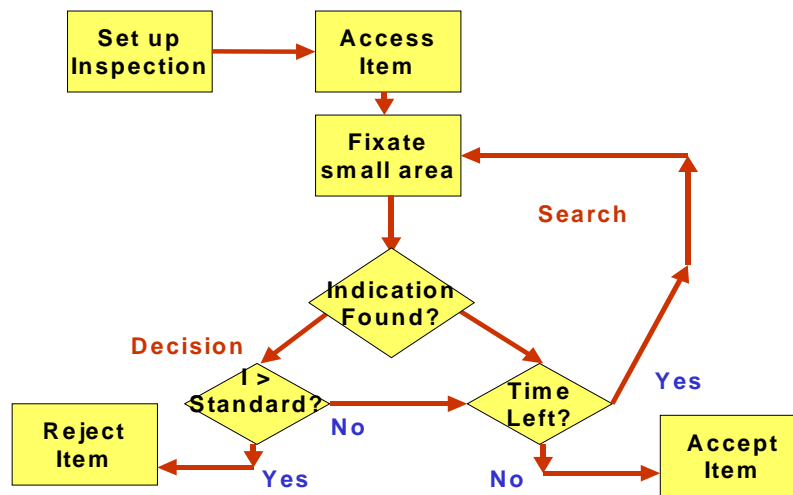


Figure 1: Model of inspection performance incorporating search and decision

Knowledge of how people perform inspection tasks in both manufacturing industries and medical diagnosis has been reviewed many times (e.g. Drury, 1997). It has also been interpreted in an aviation context following the Aloha incident (Drury, 1989; Wiener, 1989). Briefly, inspection is composed of several functions or processes, the most error-prone of which are search and decision. In search, the inspectors' eyes (or probe for NDI) move around the area to be inspected, stopping when some indication is found. In decision, this indication is compared to known (available or remembered) standards to determine whether or not a reportable fault condition exists.

A flavor of the findings of this tradition can best be given through ICAO's model of human factors in aviation: SHELL (ICAO, 1989), each element of which represents a key component of the human/ machine system. How each component interfaces with the individual considered (pilot, air traffic controller, AMT, inspector) determines the sources of both successful human performances and human errors. Table 1 summaries industrial inspection findings using this aviation-based model of human factors.

Table 1: Summary of inspection findings using ICAO's SHELL Model

SHELL System Component	Typical Findings from Inspection Studies
S: Software e.g. procedures, instructions, workcards, feedback	Instructions given to the inspector have a great effect on both p (detect) and p (false alarm). In addition, feedback information to the inspector has large positive influences on performance (Gramopadhye et al, 1997).
H: Hardware e.g. job aids, enhanced vision systems, magnifiers	Equipment such as semi-automated visual inspection systems improve performance when well-integrated with human functions (Hou et al, 1993). Enhanced vision systems, such as magnification or lighting aids sometimes help, sometimes do not. Providing visible comparison standards improves decision.
E: Environment e.g. lighting, thermal, noise	Some effects, but only at relatively extreme values and with long exposure times.
L: Liveware (Individual) e.g. individual inspector characteristics	Some general characteristics of “good” inspectors, such as field independence and peripheral visual acuity. Often each inspection task shows performance correlations with different individual characteristics.
L: Liveware/Liveware e.g. interactions with other people in system	Job design is important. Inspectors tend to feel their jobs isolate them from others. Expectations of others can have large effects on what gets reported as fault.

Applicable Knowledge 3. Human Factors in Aviation Operations:

There is a long tradition of human factors analysis of both the tasks involved in flying/guiding aircraft, and the accidents arising from these tasks. Indeed, one of the earliest human factors studies (Fitts and Jones, 1947) analyzed 460 “pilot error” accidents and found that many were induced by poor design or placement of controls and displays in the cockpits of the time. Over the succeeding years, these studies have led to great improvements in the design of cockpits, selection procedures and pilot training programs (Wiener and Nagel, 1988).

In recent years the interest, both on the flight deck and in air traffic control, has focused on the two issues of automation and interpersonal interactions. Automation studies, again both of how tasks should be performed and the accidents arising when they are performed incorrectly, have led to changes in automation systems (e.g. Phillips, 1998).

Interpersonal relations on the flight deck have also been studied both analytically and through accident analysis. From this, work has emerged a body of theory and practice known generally as Resource Management. Crew Resource Management (CRM) is now a regular, and potentially ICAO mandated, component of flight training and retraining programs (e.g. Heimreich, Foushee, Benson, and Russini, 1986; Foushee and Helmreich, 1988). Pilots (and others) learn techniques for working together more effectively from flight planning through to handling of unplanned incidents. Such results have found rational applicability in the aviation maintenance domain, now becoming known as Maintenance Resource Management (MRM). Taylor (1991) has taken a socio-technical systems approach to analysis of inter-personal activities in maintenance. This has led to training programs (e.g. Robertson, 1996; Komarniski, Russell and Johnson, 1996) which have been successful in changing attitudes and behaviors of maintenance personnel.

Using Applicable Knowledge:

Aircraft inspection has already benefited from some of these knowledge areas. Thus the ECRIRE program (Spencer and Schurman, 1995) examined one NDI technique, eddy-current inspection, incorporating human factors variables. They were able to test one-person versus two person teams (no consistent effects) and gross body posture (a small decrease in detection performance when the inspector had to work at about knee height). A FAA program on human factors in aviation maintenance and inspection (e.g. Drury, Shepherd and Johnson, 1997) has had some success in improving documentation design, lighting and communications. This program expanded the search-plus-decision model following industrial inspection findings to include five generic inspection functions (Drury, 1992):

<u>Initiate</u>	inspection, e.g. calibration, documentation
<u>Access</u>	area to be inspected, e.g. by removing access hatches
<u>Search</u>	area by successive fixations or probe movements
<u>Decision</u>	on whether indication exceeds standard
<u>Response</u>	by signing inspection as complete or recording defect.

Such a task description invites task analysis, which would lead naturally to human reliability analysis (HRA). Indeed, perhaps the earliest work in this field applied HRA techniques to construct fault trees for aircraft structural inspection (Lock and Strutt, 1985). The HRA tradition lists task steps, such as expanded versions of the generic functions above, lists possible errors for each step, then compiles performance shaping factors for each error. Such an approach was tried early in the FAA's human factors initiative (Drury, Prabhu and Gramopadhye, 1990), but was ultimately seen as difficult to use because of the sheer number of possible errors and PSF's. It is occasionally revised, e.g. in the current FRANCIE project (Haney, 1999) using a much expanded framework that incorporates inspection as one of a number of possible maintenance tasks. Other attempts have been made to apply some of the richer human error models (e.g. Reason, 1990; Hollnagel, 1997; Rouse, 1985) to inspection activities (Latorella and Drury, 1992; Prabhu and Drury, 1992; Latorella and Prabhu, 1998) to inspection tasks. These have given a broader understanding of the possible errors, but have not helped better define the PoD curve needed to ensure continuing airworthiness of the civil air fleet.

Two Recent Studies:

To help understand how human factors can contribute to the domain of aircraft inspection, two examples are given. The first pursues an analytical approach based on a task breakdown, while the second examines broader issues affecting human reliability.

The first study was the Visual Inspection Research Program (VIRP) undertaken for the FAA using a retired Boeing 737 test aircraft at Sandia National Laboratories (Drury and Spencer, 1997). Twelve experienced airline inspectors performed ten different inspection tasks, nine on the aircraft and one on a series of fuselage test panels containing known cracks. The total experiment lasted 1.5 to 2 days per inspector and was performed under highly realistic conditions in a flight hangar. Overall, performance was quite variable. Inspectors took from 7.5 to 12.3 hours of inspection time for the ten tasks. On a set of large cracks and corrosion defects which the manufacturers would expect inspectors to find, the probability of detection was also quite variable. PoD ranged from 0.5 to 1.0 on large cracks and from 0.3 to 0.6 on large corrosion areas. There was little evidence of a speed/accuracy tradeoff across inspectors. There were also low correlations between inspector performance on the 10 tasks, and also between pre-test measures and task performance. Individual differences were large and inconsistent.

A more detailed analysis of this data is possible by classifying errors into search errors and decision errors. Drury and Sinclair (1983) showed that this was possible in industrial inspection of aircraft bearings. For the panel inspection task, we used video tape records to classify the errors. It was possible to see from the video tape whether an inspector passed quickly over a crack defect (search error) or whether he paused to examine the defect more closely before either reporting it or moving on. This latter was a decision error, either a miss or as false alarm.

Figure 2 shows the individual differences between inspectors for search performance. Note that probability of search success is rather uniform and low. The mean was 0.5 and the coefficient of variation was 0.2. For decision performance, Figure 2 shows individual inspector results plotted on Relative Operating Characteristic space. The variability is readily apparent, with mean performance as follows:

p (correct hit): mean = 0.84 CV = 1.2
 p (correct No): mean = 0.64 CV = 1.0

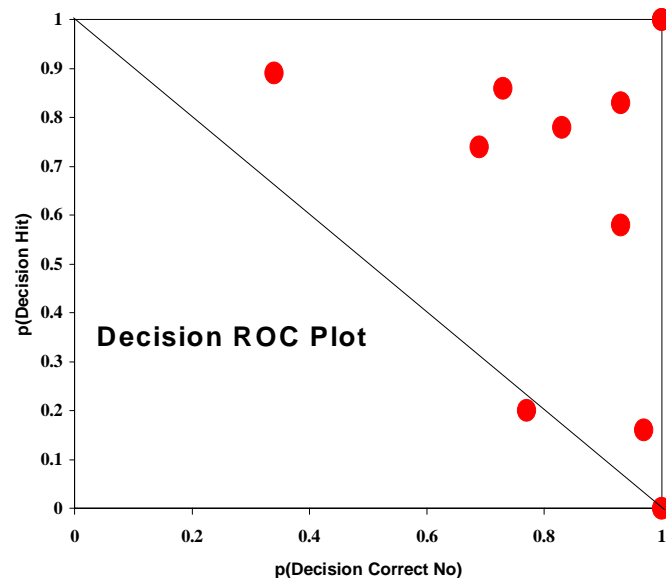


Figure 2: Relative Operating Characteristic (ROC) curve for decision component in aircraft structural inspection. Each point is from one inspector.

Thus, search performance could be characterized as consistently poor, whereas decision performance was better, but highly variable. Search and decision performance were statistically unrelated. Such findings allow us to focus interventions, for example by improving lighting and training to support search, while using training and feedback to reduce inter-inspector variability in decision (Gramopadhye, Drury and Prabhu, 1993).

The second experiment was similar to VIRP, but performed on a commuter aircraft (Fairchild Metro) using experienced regional airline inspectors. This study had inspectors perform seven inspection tasks, again over 1.5 to 2 days. Performance was again highly variable between inspectors and between tasks with almost no correlations between task performance or between task performance and pre-tests. However, in this experiment, the major concentration was on the subtleties of the inspection task and its context, detailed in Wenner and Drury (1997).

References

- Biederman, I. and Shiffar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual learning task. *Journal of Experimental Psychology, Learning, Memory and Cognition*, **13**(4), 640-645.
- Bobo, S. (1989). Communication and transfer of non-destructive inspection information. In *Human Factors Issues in Aircraft Maintenance and Inspection - Information Exchange and Communications*, Second Federal Aviation Administration Meeting. Washington D.C.
- Boeing, (1997). *Statistical Summary of Commercial Jet Airplane Accidents*, Boeing Commercial Airplane Group, Seattle, WA.

- Chi, C.-F. and Drury, C. G. (1998). Do people choose an optimal response criterion in an inspection task? *IIE Transactions* (1998), **30**, 257-266.
- Courtney, A. J. (1984). A search task to assess visual lobe size, *Human Factors*, **26(3)**, *Human Factors*, 26 (3), 289-298.
- Dalton, J. M. (1991). *Visual search in inspection of sheet steel*, Unpublished M.S.Thesis, State University of New York at Buffalo.
- Drury, C. G. (1975). Inspection of Sheet Materials - Model and Data. *Human Factors*, (17), 257-265.
- Drury, C. G.(1989). Industrial Inspection, Proceedings of the Second International Conference in Ageing Aircraft, Baltimore, 213-220.
- Drury, C. G. (1990). Visual search in industrial inspection. In D. Brogan (Ed.), *Visual Search*, London, UK, Taylor & Francis, Ltd, 263-276.
- Drury, C. G. (1992). Inspection Performance. In G. Salvendy (Ed.), *Handbook of Industrial Engineering*. New York, John Wiley & Sons, 2282-2314.
- Drury, C.G. (1994). The speed-accuracy trade-off in industry. *Ergonomics*, **37**, 747-763.
- Drury, C. G. (1997). Human Factors Audit, Chapter 48. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics*, New York, John Wiley & Sons, 1593-1616.
- Drury, C. G. and Addison, J. L. (1973). An industrial study of the effects of feedback and fault density on inspection performance. *Ergonomics* (16), 159-169.
- Drury, C. G. and Sinclair, M. A. (1983). Human and machine performance in an inspection task. *Human Factors*, **25**(4), 391-400).
- Drury, C. G. and Spencer, F. W. (1997). Human factors and the reliability of airframe visual inspection. *Proceedings of the 1997 SAE Airframe/Engine Maintenance & Repair Conference* (AEMR'97, August 1997).
- Drury, C. G., Prabhu, P. and Gramopadhye, A. (1990). Task Analysis of Aircraft Inspection Activities: Methods and Findings, *Proceedings of the Human Factors Society 34th Annual Conference*, 1181-1185.
- Drury, C. G., Shepherd, W. and Johnson, W. (1997). Error reduction in aviation maintenance, *Proceedings of the 13th Triennial Congress of the International Ergonomics Association*'97, Tampere, **3**, 31-33.
- Federal Aviation Administration (1993). *National Aging Aircraft Research Program Plan*, Atlantic City, NJ, FAA Technical Center .
- Fitts, P. M. and Jones, R. E. (1947). Analysis of factors contributing to 460 "pilot error" experiences in operating aircraft controls (Memorandum Report TSEA-4-694-12, Aero Medical Laboratory). *Selected papers on human factors in the design and use of control systems*. H. W. Sinaiko. Wright-Patterson AFB, Dover, Harry Armstrong Aerospace Medical Research Laboratory.
- Foushee, H. C. and Helmreich, R. L. (1988). *Group interaction and flight crew performance*. Chapter 13, In Wiener, E. L. and Nagel, D. C. (Eds.) *Human Factors in Aviation*. San Diego, Academic Press, 189-225.
- Gramopadhye, A., Drury, C. G. and Prabhu, P. V. (1993). Training for Visual Inspection of Aircraft Structures. *Human Factors in Aviation Maintenance, Phase 3: Progress Report*, Interim Report, DOT/FAA/AM-93/15, Springfield, VA, National Technical Information Service.

- Gramopadhye, A. K., Drury, C. G. and Sharit, J. (1997). Feedback strategies for visual search in airframe structural inspection. *Int. Journal of Industrial Ergonomics*, **19**(5), 333-344.
- Goranson, U. F. and Rogers, J. T. (1983). Elements of Damage Tolerance Verification,. *12th Symposium of International Commercial Aeronautical Fatigue*. Toulouse.
- Haney, L. (1999). FRamework Assessing Notorious Contributing Influences for Error (FRANCIE), *Proc. International Workshop on Human Factors in Space*, Tokyo, July 1999.
- Harris, D. H. (1969). The nature of industrial inspection. *Human Factors*, **11**(2), 139-148.
- Heimreick, R. I., Foushee, H. C., Benson, R, and Russini, R. (1986). Cockpit management attitudes: Exploring the attitude-performance linkage. *Aviation, Space and Environmental Medicine*, **57**, 1198-1200.
- Hollnagel, E. (1997). CREAM-Cognitive Reliability and Error Analysis Method. New York, Elsevier Science.
- Hou, T.-S., Lin, L. and Drury, C. G. (1993). An empirical study of hybrid inspection systems and allocation of inspection function. *International Journal of Human Factors in Manufacturing*, **3**, 351-367.
- Hovey, P. W. and Berens, A. P. (1988). Statistical evaluation of NDE reliability in the aerospace industry. In D. D. Thompson and D. E. Chimenti, *Review of Progress in Quantitative Nondestructive Evaluation*. New York, Plenum Press, **7B**, 1761-1768. .
- ICAO (1989). *Human Factors Digest No. 1 Fundamental Human Factors Concepts*, Circular 216-AN/131, Canada, International Civil Aviation Organization.
- Komarniski, R., Russell, B. and Johnson, W. B. (1996). Perspectives on TRM training for maintenance. *FAA/NASA Tenth Meeting on Human Factors Issues on Aviation Maintenance and Inspection (January 17-18)*. Alexandria.
- Kundel, H. L. (1975). Peripheral vision, structured noise and film reader error. *Radiology*, **114**, 269-273.
- Latorella, K. and Drury, C. G. (1992). A framework for human reliability in aircraft inspection. In *Proceedings of the 7th FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, Washington, D.C., Federal Aviation Administration.
- Latorella, K. and Prabhu, P. (1998) Human error in aviation maintenance and inspection, *International Journal of Industrial Engineering*
- Lock, M. W. B. and Strutt, J. E. (1985). Reliability in in-service inspection of transport aircraft structures, CAA Report 95013, London, Civil Aviation Authority.
- Megaw, E. D., Alexander, C. J. and Richardson, J. (1979). Fault mix and inspection performance, *Int. J. Prod. Res.*, **17**(3), 181-191.
- Murgatroyd, R. A., Worrall, G. M., Drury, C. G. and Spencer, F. W. (1997). Comparison and Further Analysis of CAA and FAA inspection Reliability Experiments, CAA Paper 96010, London.
- Murgatroyd, R. A., Worrall, G. M. and Waites, C. (1994). *A Study of the Human Factors Influencing the Reliability of Aircraft Inspection*, AEA/TSD/0173, U.K., AEA Technology.
- Phillips, E. H. (1998). Inspection Methods "Key" to Aging Aircraft Safety. *Aviation Week*, **70**, 719.
- Prabhu, P. and Drury, C. G. (1992). A framework for the design of the aircraft inspection information environment. In *Proceedings of the 7th FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, Washington, D.C., Federal Aviation Administration.

- Reason, J. (1990). *Human Error*, Cambridge, Cambridge University Press.
- Robertson, M. (1996). Team training in aviation maintenance operations. In O. Brown Jr. and H. W. Hendrick (Eds.), *Human Factors in Organizational Design and Management*, The Netherlands, Elsevier Science, 583-588.
- Rouse, W. B. (1985). Optimal allocation of system development resources and/or tolerate human error, *IEEE Transactions: Systems, Man and Cybernetics*, SMC-15, **5**, 620-630.
- Rummel, W. D., Hardy, G. L. and Cooper, T. D. (1989). Applications of NDE reliability to systems. *Metals Handbook*, **17**, 674-688.
- Spencer, F. and Schurman, D. (1995). Reliability Assessment at Airline Inspection Facilities. Volume III: Results of an Eddy Current Inspection Reliability Experiment, DOT/FAA/ CT-92/12. Atlantic City, FAA Technical Center.
- Taylor, J. C. (1991). Maintenance organization. *Human Factors in Aviation Maintenance Phase 1: Progress Report*, DOT/FAA/AM-91/16, Chapter 2, Office of Aviation Medicine, Springfield, VA, National Technical Information Service, 15-43.
- Thomas, L. F. and Seaborne, A. E. M. (1961). The sociotechnical context of industrial inspection, *Occupational Psychology*, **35**, 36-43.
- Wenner, C. (1999). *The Role of Instructions in the Aircraft Maintenance Domain*. Unpublished Master's Thesis, Buffalo, NY, State University of New York at Buffalo.
- Wenner, C. and Drury, C. G. (1997). Human error and ground damage. *Proceedings of the 9th International Symposium on Aviation Psychology*, Columbus.
- White House Commission on Aviation Safety and Security (1997). *Gore Commission Report*.
- Wiener, E. L. (1989). The vigilance phenomenon. *Proceedings of the Second International Conference in Ageing Aircraft*. Baltimore, 206-212.
- Wiener, E. L. and Nagel, D. C. (1988). *Human Factors in Aviation*. San Diego, Academic Press, Inc.

Effects of Practice and Memory Aiding on Decision Performance and Information Search in Command and Control

Peter H.M.P. Roelofsma
 Vrije Universiteit
 Vakgroep Psychonomie Provisorium
 De Boelelaan 1111
 1081 HV, Amsterdam
 The Netherlands

Abstract

This study examines to what extent practice (learning by doing) and memory aiding (paper and pencil) affects human decision performance and information search in command & control. A management command control task which simulates a dynamic competitive environment was used for this purpose. In six practice sessions memory aiding was systematically varied. 104 subjects participated in the simulation task. Their success score, failure score and decision speed score (number of decisions made per session) were measured as well as their information search profile. The results show that memory aiding led to a decrease of the number of decisions made and eventually resulted in a decrease in the overall success score. It led also to a decrease in the amount of information that subjects searched. Subjects used a satisficing decision making strategy, using little information, which proved however to be successful in the long run. Practice resulted in a slight but significant increase in the amount of information search. However, most information was searched after the decision as a means of justifying the decision. It is concluded that memory aiding makes subjects spend too much time in problem structuring, at least relative to the dynamics of a complex command and control environment. The results are discussed in terms of cognitive continuum theory.

1 Introduction

A major issue in the performance of complex man-machine systems concerns the human decision processes in such systems. Complex man-machine systems share at least two essential features. First, the amount of information flow that is used for the decision making process is very large and the variables representing the information are directly or indirectly connected by a relational network which is often only partially known to the decision maker. Second, the decisions in these systems have to be made within a restricted amount of time. As a result human operators are often put in a situation of time pressure. High information flow and time pressure are two main characteristics of current complex man-machine systems, examples are command and control (C2) in military units, fire fighting units, medical teams and business dealing rooms.

Information load and time pressure make the decision process in C2 a task beyond the ability of a single individual. Therefore information has to be distributed over several individuals. The crucial question then becomes how an optimal distribution of information can be achieved so that efficiency and effectiveness in the decision making process is maximised. It is in this context of extreme importance for C2 design to assess what information is indeed relevant for the decision making process and what information can be eliminated. At the heart of these issue lies the crucial question that should be answered first: how much information can the decision-maker attend to when he or she is under pressure?

Almost 15 years ago, Andriole & Hopple (1986) proposed that issues of human factors, like the decision processes underlying C2, have been ignored by engineers and that the design and development of C2 systems has been dominated mainly by interest in technological advances. At the turn of the millennium, awareness for human factor issues has generally increased by designers and planners of C2. This realization of the necessity of human factors resulted in an increasing need for applied psychological research issues, among others the before mentioned question concerning how much information the decision-maker actually can deal with. This report focuses on this last issue: How much information do people use in assessing the value of a prospect in

C2? How is decision performance influenced by practice and memory aiding? Is the information they search for used in an efficient way?

In the psychological literature the early studies of Einhorn (1974) and Ebbesen & Konecni (1975) have already shown that even expert decision-makers use only relatively little information in their judgements. It should be mentioned that these studies deal essentially with static, routine decisions without time pressure. When these results are generalised to C2 situations this may portray an unrealistic view. Still, the conclusions are striking enough. For instance, Einhorn (1974) showed that medical doctors use no more than three symptoms in their diagnostic process. Ebbesen and Konecni (1975), in a study on court psychology showed that judges did use not more than one to three aspects in their decision making process on sentencing defendants. Gigerenzer & Goldstein (1996) mentioned examples of experts using mainly one essential cue. Following Simon (1976), Gigerenzer & Goldstein's basic claim is that decision-makers are satisficing. That is, all the decision-maker does is considering alternative courses of action sequentially until the first cue that 'will do' is found. As a potential explanation Geath and Shanteau (1984) suggested that many decision-makers have difficulties in distinguishing relevant from irrelevant information. Since irrelevant information needs attention as well, it becomes a heavy burden in information processing. Indeed, this may account for the earlier findings of limited use of information. This means that people use not only a limited amount of information, but also that irrelevant information may have a negative effect on the amount of relevant information that is eventually evaluated.

A related issue is that people seem to have an intrinsic tendency to search for redundant and irrelevant information. For instance, Lusted (1977) found that radiologists use information inefficiently. They search for additional information even when it is redundant and non-diagnostic. A likely explanation for this finding is that the information is not used during the decision making process itself, but that it is mainly applied for justifying the decision. That is, subjects do not use the information before the decision is made, but it is used afterwards, as a means of defending the decisions they have already made. This hypothesis is in line with Festinger's (1957) notion that decision making is mainly a search for justification. Festinger's work recently has gained a revival in the decision-making literature on post-decisional processes like regret and uncertainty reduction (Svenson & Benthorn, 1992).

An important question in this context is: What happens when the subject receives practice, e.g. learning by doing, and memory aiding during the decision task? Will subjects be able to search for more or different types of information and will decisional performance improve as a result? Or will subjects search only for more justification? This question is important not only from a theoretical point of view but perhaps even more for practical reasons. For example, the development of training programmes of C2 decision making is becoming increasingly popular. But the popularity of training programmes stems mainly from common sense notions like 'practice makes perfect'. In addition, most aiding devices that are planned and developed are often based on similar common sense notions. For example, memory aiding is supposed to reduce the workload, to support the representation of the environment and thereby improve decision performance. However, basing system design only on such notions is like relying more on belief than on facts, since there is hardly any evidence that in complex decision making under time pressure either practice or aiding will help. The results of the few studies available do not generally support the hypothesis that mere practice and aiding are beneficial. For instance, Broadbent, Berry and Gardner (1990) did not find any practice effects on decision performance in a complex command and control task.

There are several ways how practice and aiding can influence decision performance and the amount of information that subjects use. Take the following general scenario as an example. Suppose a decision-maker is confronted with a situation in which a series of prospects pass the scene. For instance, a series of potential threat alarms in military command and control. Or a series of buy options in a business dealing room. For each of these prospects the decision-maker can search for information about the value of the prospects, i.e. the seriousness of the threat or the (un)attractiveness of the offer in a number of ways. The decision-maker may search **locally** which of the prospects in one particular location or place is the most serious, and adjust the limited resources accordingly. Or the decision-maker may search more **globally**, that is over several locations, and assess the range of potential (un)attractive buyoffers or threats over different places. Finally, the decision-maker may assess the value of a prospect by taking information concerning **future** developments of the

system into account. A threat may represent a forebode of a potential larger threat, like the outbreak a war or the total collapse of the economy. So, the decision-maker can assess the value of a prospect by looking for more information locally, globally and for future developments of the system.

One approach towards learning and practice in complex dynamic decision making holds that suboptimality in decision performance is the result of insufficient knowledge of the decision environment or insufficient capacity to store information about the decision environment. As a consequence subjects are *forced* to rely on primitive heuristics (see: Kleinmuntz, 1985; Hogarth, 1981; 1987; Wickens, 1992). If subjects were able to improve their knowledge representation by searching for more information about the decision environment, e.g. by practice or by using memory aids, they would rely less on such primitive decision principles. They will gradually improve their decision performance by combining and integrating more information about alternatives and about threat dimensions and use this before the decision is made.

Table 1.1: Transfer table showing systematic variation of aiding and practice

No Aid	Aid
Aid	No Aid
Aid	Aid
No Aid	No Aid

This issue can be addressed by studying the transfer effects of a memory aiding manipulation. A transfer table is depicted in Table 1.1. which represents how memory aiding can be systematically varied with practice. Table 1.1. gives insight in the potential relationships between aiding, practice and decision performance. In particular, if memory aiding leads to the use of more information and to a better knowledge representation the following four questions are relevant: (1) To what extent does memory aiding improve knowledge acquisition? This question can be answered by comparing the performance of a group that does not receive any memory aiding (no aid –no aid) with a second group that starts with aiding which is removed in the second part of the experiment (aid-no aid). In the case of positive transfer memory aiding has promoted knowledge acquisition. If in the second group performance drops during the second part of the experiment to that of the first group, memory aiding has been useful but has not promoted acquisition of knowledge representation. There could be even negative transfer of memory aiding resulting from merely relying on aiding rather than acquiring a better representation. (2) Is memory aiding just a matter of extending working memory? This question can be answered by comparing the decision performance of a group that starts without aiding and receives aiding in the second part of the experiment (no aid-aid) with a second group that starts with aiding in the first part, and receives no aiding in the second part (aid-no aid). If memory aiding is only a matter of supporting the working memory the patterns of results for these two experimental groups should mirror each other. (3) To what extent is a successful decision performance possible without any memory aiding? This question can be answered by comparing a group that receives aiding (aid-aid) with a group that does not receive any aiding (no aid – no aid). (4) Is memory aiding only helpful after sufficient practice? This question can be answered by comparing a group that does not receive any aiding (no aid – no aid) with a group that receives no aiding in the first part of the experiment but only in the second part of the experiment (no aid – aid).

There is an alternative line of argument that stems from the notion that decision making is these complex dynamic decision making is best performed by intuitive cognition (Hammond et al, 1987). Hammond et al argued that cognitive tasks can be classified on a cognitive continuum. Some tasks are intuitive inducing others are analysis inducing. Some properties on intuition inducing tasks are: the number of cues is larger than 5 and there is a high redundancy and simultaneous display of cues. In addition there is a low certainty level in the task and the decision time period is brief. These are aspects which are all prototypical for C2 situations. C2 tasks are therefore prototypical intuition inducing tasks.

Intuitive cognition in decision making is guided by inbuilt satisficing principles (Simon, 1957). Following a satisficing principle subjects base their decision on the **minimal** amount of information needed to come to a decision. They focus mainly on the current and imminent options. As mentioned earlier, satisficing means that the decision maker mainly considers alternative courses of action sequentially until the first cue that

satisfies a subjectively predetermined criterion is found. Subjects will take 'the first and the best' option. They neglect alternative options and evaluate on the basis of mainly one dimension or aspect, which may vary from context to context. It is obvious that this does not necessarily always lead to the optimal outcome, but the strategy is very flexible and may save much processing time. Indeed, the strategy may even prove to be highly successful in complex and dynamic situations and it may even outperform an analytical decision making strategy (Hammond et al, 1987; Gigerenzer and Goldstein, 1996).

Performance improvement will be the result of a more differentiated value system, e.g. due to practice. This may result in changes in the minimal criteria chosen to evaluate an option as appropriate or not. In general, there is no expectation of an increase in the total number of alternatives and dimensions that are evaluated before the decision is made. The decision making process, even for experts, remains basically a process of satisfying principles, and it is even more likely that experts will use less information. For example, novices who try to find out what the causes are for a sudden dysfunction in their car appear to search for much more information than an expert repairman who may only need one or two cues.

Following this line of thinking one can be very sceptic about the effects of memory aiding that aim at promoting information processing, since it may slow down the decision making process. Thus, storing and saving more information for the decision making process may increase the decision processing time and decrease decision flexibility in a dynamic environment. As a result of memory aiding subjects may consider more analytical judgement processes in the sense that subject may store more information which they may try to integrate in the decision making process. Consequently subjects decision performance can even be diminished as a result of memory aiding.

The concept of satisficing has been demonstrated in numerous experimental tasks in the area of behavioral decision making (e.g., Simon, 1955; Huber *et al.*, 1990; Dörner, 1987; Gigerenzer & Goldstein, 1996). However, these demonstrations were solely based on static decision tasks. The concept of satisficing has received little attention in the literature of complex dynamic decision making. In particular, the effects of practice (learning by doing) and memory aiding have received little attention. In fact, albeit the long history of complex and dynamic decision making, experimental studies conducted in this area are still relatively scant in general (Brehmer, 1987; Flin et al 1997; Kerstholt, 1996; Mynatt et al 1977; Roelofsma, 1995; Zsombok & Klein, 1997). In order to study the potential negative or positive effects of practice and memory aiding on decision performance and information search in a simulated C2 task, in the present study memory aiding will be systematically varied over a series of practice sessions.

There are several problems in regard to experimental studies with simulated C2 tasks. First, the task must capture the major characteristics of a generic C2 task, such as:

- a) it should have an objective that is both well defined but for which there is also no simple strategy or solution.
- b) performing the task should require analyzing several sources of information, that differ in aspects like timeline or modality.
- c) the task should have a judgment component as well as a choice component. Subjects must assess a situation, evaluate it and diagnose it in order to make decisions.
- d) the decisions in the task should be made within a restricted amount of time.

A second problem in studying aspects of practice and aiding is that the experimental task should neither be too simple nor too complex and there should be sufficient time for the subjects to acquire expertise in the task. Indeed, some of these problems may be the reason why Broadbent et al. (1990) did not find any improvement in decision performance as a result of practice.

A third problem is that the task should generate sufficient performance measures. For example, the task should generate a meaningful 'success' score. In military command and control this may refer, for instance, to the number of successfully eliminated targets per unit of time. In dealing room C2, this could be expressed in the profit made per decision. Next, there should be an error or failure score, such as the amount of losses in military C2 or the number of bankruptcies in business C2. But it also should give a measure of decision speed, such as the total number of decisions in a fixed amount of time. The combination of the above measures is

especially important for interpreting decisional performance in C2, since the decision maker may, for example, improve his success score, while at the same time the costs of the losses or the speed of the decision process become unacceptable.

Finally, in order to examine cognitive aspects like information search the task should generate a profile of the information that the decision maker uses before and after the decision. For example, information about the value of current and imminent prospects, as well as information about local, global and future values. A C2 task that meets these criteria is MILSIM developed by Van Doorne (1986) and this task was used in this experiment.

2 Method

Subjects. Hundred and four VU undergraduate psychology students applied to participate in this experiment in turn for study credits, equivalent to 40 hours study time. Subjects were aged between 18 and 23 years.

The C2-task. Subjects make trading decisions in a business command control simulation of a complex and dynamic trading environment. The game simulates a competition between several merchant fleets, each of which owns several shuttles and all of which are competing for trading opportunities. The headquarters of the fleet constitute a C2 system in which the market situation is continuously assessed and decisions made concerning the portfolio (or consumption bundle) of the products in a spaceship. The goal of a decision maker (DM) in MILSIM is to maximize profit by buying products at centers that offer low prices and sell them at centers that offer higher prices. There are six trading centers and six commodities that can be traded. Each center sells three commodities and buys three commodities. A center does not buy and sell the same products. The demand and supply in terms of quantities as well as prices differ from one center to the other and are continuously changing as a result of the economic scenario's that evolve during the game. For that purpose constant information has to be obtained on the supply and demand of different commodities in each center and about the future developments of the environment. If the DM makes the right decisions the company will prosper and make large profits. Alternatively, wrong decisions may lead to losses and eventually to bankruptcy.

The structure of the game is as follows. The DM sits behind a computer screen. Then the DM receives information that a specific ship has landed on a certain centre in space. Next, the DM starts a so-called 'dialogue' with a simulated captain of the ship. We will refer to this action as SD (start dialogue). In this dialogue the DM will be sequentially confronted with three sell offers and three buy offers (prospects). That is one offer at a time. We will refer to the sell offers as S1, S2 and S3 and to the buy offers B1, B2, and B3. For each of the buy and sell modes the following information is presented to the DM on the dialogue screen. Some of the information can be crucial for a decision; other information is simply irrelevant. The specific contents of the information presented below is used only here to illustrate the procedures and this is printed in *Italics*:

- 1) The message containing the buy offer e.g.: 'Centre sells food'.
- 2) The quantity of the offer, e.g. '24 kugels'. (kugels is a fictitious unit for amount)
- 3) The price per kugel, e.g. '2000 / kugel'
- 4) The name of the centre location, e.g. 'Alcor'
- 5) The number of the center, e.g. '#cntr 2'
- 6) The name of the ship, e.g. '*Our Mary*'
- 7) The ship number, e.g. '#sh 3'
- 8) The name of the captain of a ship, e.g. '*Peter*'
- 9) The current date, e.g. 'Day:14 Month: 4 Year 2525'
- 10) Visiting time on center, e.g. 'Minutes: 4 Seconds: 0'
- 11) The price of info, e.g. '800'
- 12) The ships money, e.g. '80 000'
- 13) The ships weight, e.g. '25'
- 14) Travel costs, e.g. '4000' per 100 starmiles
- 15) The commodities that the ship has on board. A table is presented on the screen which tells which product are stored on board and how many.
- 16) The weights of the products. This is also presented in a table.

- 17) The distances between the starcenters. This is presented in a star map.
- 18) The products that were bought on the current center. This is presented by two asteriks (**) in the commodity table.
- 19) The fuel price, e.g. '1.8'.

The travel costs are a function of the ships weight, the travel distance and the fuel price. The fuel price fluctuates, as well as the information prices. The ships travel time is a function of the weight of the ship, the distance and fluctuations in the environment. For a description of the simulation model see table 2.1 in appendix II and Van Doorne (1986). Obviously, the DM is not aware of precise nature of the model.

Table 2.1: Overview of the experimental design

	3 * 1½ hour sessions	3 * 1½ hour sessions
Group I (n = 26)	No Memory aiding	No Memory aiding
Group II (n = 26)	Memory aiding	No Memory aiding
Group III (n = 26)	No Memory aiding	Memory aiding
Group IV (n = 26)	Memory aiding	Memory aiding

As can be seen in the list above the DM receives much information, but there is one aspect on which the DM is always left in uncertainty. This concerns the uncertainty related to the relative value of the current prospect, e.g. what is the value of 24 kugels of FOOD? To assess the relative value of the current option in B1 the DM can search for several pieces of information. First the DM can search for local alternatives. That is, the DM can purchase information on the alternative options of the center where the ship is currently located. This we have labelled **local** info. If the DM buys this info he receives the demand and supply of the current center in terms of prices and amounts. Alternatively, the DM can search for global alternatives. That is, he may search for information about the values of other products at other centers. We have labelled this type of information **global** info. Then the DM may search for the future value of products. We have labelled this info: **future** info. When the DM buys this info, information is presented on the future fluctuation of prices and quantities in the simulated economy, as well as development of fuel and gas prices and general travel times, potentials wars and disasters. The DM can also buy only the info concerning the value of the current offer under consideration. We have labelled this info **imminent** info. Buying imminent info about the product FOOD, for example, means that the DM receives info on the current demand and supply of FOOD in terms of buy and sell prices and the corresponding quantities. Of course, the DM may always decide not to buy any information at all and trust only on his intuition about whether or not to accept the offer. The DM has to indicate about how much will be bought -if any- by entering the corresponding value. Entering the value 'zero' means that the DM refuses the offer. Then the DM automatically arrives in B2. The second prospect is offered and the whole procedure described under B1 is repeated. Next follows B3. The DM has in total 4 minutes to decide about his consumption bundle for B1-B3 together. When these 4 minutes have passed, or after the DM has finished with B3, the DM automatically comes in the so called travel destination mode (TD). Here the DM has to decide about the new travel destination for the ship. This decision has to be made within 30 seconds. When this time mode is surpassed the ship will be charged with 2500. If the ship has insufficient money to pay for the travel costs the center will buy back the products that the DM had just bought. But now the currency will be only half of the price that the DM had just paid for. In the TD the DM can again buy each of the information types described above. We will refer to this type of information search: **after**. When the ship has insufficient money to make one more travel the company is bankrupt and the DM has to start all over again.

The ship starts with an empty ship and a capital of 80 000. On the arrival at the new location the DM first passes the three sell modes. Then the Buy modes sequence starts all over again. We will refer to a complete decision round to a center as a **visit**.

Procedure. Subjects were tested in six sessions of 1½ hour. The sessions were divided over one week. Apart from these sessions there was a general instruction session of 45 minutes. In this general instruction session

the subjects read the general game scenario instructions (see appendix I). Next they learned how to play the game at a keystroke level. After this first general practice they were required to re-read the instructions for about 5-10 minutes. Six subjects were tested at a time in this experiment. They were visually separated from each other, although all subjects were in one large experimental room. The best out of each six players received a HEMA Dutch apple pie with a value of Fl 9.95.

Design. Subjects were randomly divided into four groups as is shown in Table 3.1. As mentioned earlier, these four groups are required to deal with the simulation game in six experimental sessions of one and a half-hour. The use of memory aiding was systematically varied within and between subjects (see Table 3.1) in a standard transfer design. In the memory aiding condition subjects received paper and pencil to support the decision making and they were told to use it as a support for their decision making process. In the no aid condition no paper and pen were available for the subjects. Table 3.1 shows that subjects can start either with or without aiding and end with or without aiding. We will refer to this as the **begin aid** manipulation and **end aid** manipulation. The four groups we will refer to as **no aid-no aid**, **no aid-aid**, **aid-no aid** and **aid-aid**. By comparing the main and interaction effects of the begin aid and the end-aid manipulation with practice the potential positive or negative effects of aiding can be examined.

For each subject, the amount of visits in one session was measured, as well as the number of bankruptcies per session, the profit made per visit, the amount of info per visit. A MANOVA was used to analyse these results. The information types: imminent, local, global, future and no info were measured and analysed descriptively over all groups.

3 Results

3.1 Profit making. In order to give an illustration of subjects' profit making behaviour a sequence chart is plotted in figure 3.1 for one of the subjects. The chart depicts the subject's profit for each subsequent visit. It can be compared with the 'heartbeat' of the profit making process. A reference line is printed at the mean of the series. For the particular subject in the figure it shows that profit gradually goes up with an occasional extreme peak down- or upward. These extremes are a potential problem for further statistical analysis when the mean is used as the measure for central tendency in techniques that require a normal distribution. The procedure that was followed in this study was to identify and exclude extremes relative to the distribution of the 'profit' variable over all sessions and groups. Explorative statistics with stem and leaf plots was used to achieve and justify this.

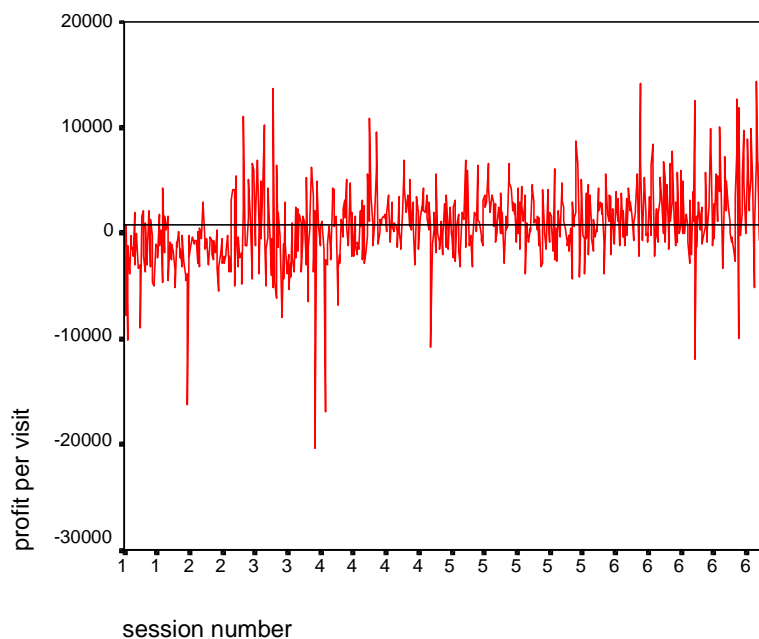


Figure 3.1: Profit Sequence Chart for subject no 16

In figure 3.2. and table 3.1 (see appendix III) the mean profit per visit is presented for each group. A strong increase in the mean profit over sessions can be observed for all groups. Subjects start with loosing about 1500 per visit and over sessions the profit per visit increases up to about gaining 1750 per visit. Indeed, this practice effect is highly significant ($F=54.061$, $df = 96$, $p < .0001$). There is no interaction effect of the begin- and end aiding ($F=1.40$, $df = 5$, $p=.222$) nor are there main effects of either the begin- or end aiding conditions ($F = 2,73$, $df=1$, $p = .101$ and $F = .048$, $p = .827$).

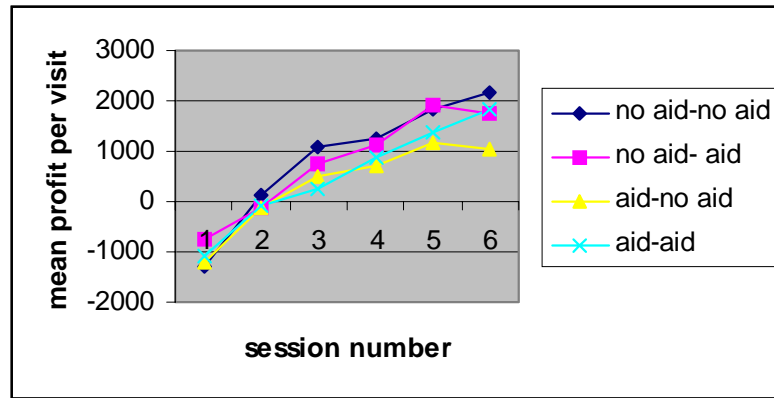


Figure 3.2: Overview of the subjects success score: mean profit per visit for each session and experimental group

Also, the interactions of practice and the aiding conditions are not significant ($F=.641$, $df = 5$, $p=.669$ and $F=1.117$, $df = 5$, $p=.357$). Finally, the second order interaction between practice and the begin and end aiding condition is not significant ($F=2,168$, $df = 5$, $p=.064$). This means that in the aiding conditions the mean profit per visit is not higher than in the non-aiding conditions. If any: aid has a negative effect on performance in the subsequent non-aid session (negative transfer).

As mentioned there is a strong effect of practice on the mean profit per visit. The within-subjects (repeated) contrasts give some further insight in the differences between subsequent sessions. It appears that the difference between session 1 and 2 is highly significant ($F=72.55$, $df=1$, $p<.001$) as well as between session 2 and 3 ($F=21.84$, $df=1$, $p<.001$) session 3 and 4 ($F=7.9$, $df=1$, $p=.006$) and session 4 and 5 ($F=14.47$, $df=1$, $p<.001$). The difference between session 5 and 6 is not significant ($F=1.20$, $df = 1$, $p=.275$). This means that in each subsequent practice session the mean profit per visit increases and which levels off only in the last session.

3.2 Bankruptcies. As mentioned subjects become bankrupt when the captain of a ship has insufficient money to make even one business travel. Figure 3.3 depicts the mean number of bankruptcies for the six practice sessions and for each of the experimental groups. As can be seen in the figure almost everyone becomes bankrupt in the first session. This clearly diminishes with practice and in the last session most subjects are able to survive throughout the session. The figure shows that in the last practice session the mean bankruptcies have dropped for all groups. A similar pattern is reflected in the median. The (main) effect of practice turned out to be significant ($F=10.01$, $DF=96$, $p<.001$). This is mainly due to a significant decline between session 1 and 2 ($F=32.77$, $DF=1$, $p<.001$). The increase from session three to four is slightly significant ($F=4.12$, $DF=1$, $p=.045$).

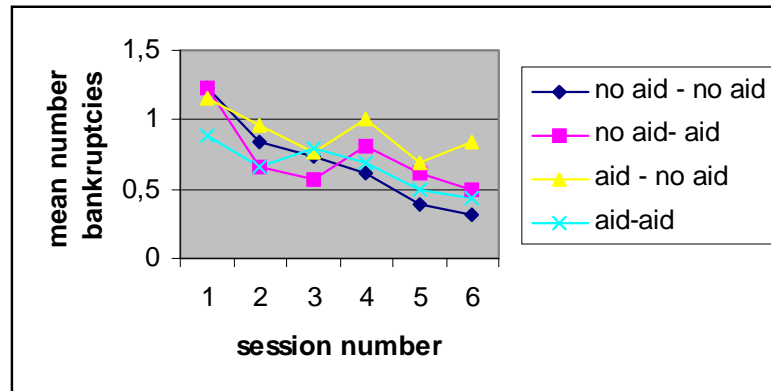


Figure 3.3: Overview of the subject's failure scores: mean number of bankruptcies per session for each of the experimental groups

3.3. Number of visits per session. As mentioned earlier, the speed of the decision making process is reflected in the number of visits that the subject can make in one experimental session. Figure 3.4 and Table 3.2 (see appendix IV) depict the total number of visits for the four groups in each of the six sessions. The figure shows two points: 1) the overall number of visits increases with practice in all conditions. Subjects start with about 60 visits per 1 ½ hour session. This doubles to about 120 visits per session. Indeed, the practice effect is highly significant ($F=45.80$, $df=96$ $p<.001$). 2) Aiding conditions show fewer visits as compared to no aiding conditions, somewhere between 15-30 visits per session. The figure shows a cross-over interaction between the aid-no aid and no aid-aid group. This effect appears to be highly significant ($F= 13.57$, $df=1$, $p<.001$). There is a main effect of the begin aid manipulation ($F=6.13$, $df= 1$, $p= .015$), an interaction effect of the begin-aid manipulation with practice ($F=7.09$, $df=5$, $p<.001$) and an interaction effect of the end aid effect with practice ($F=3.635$, $df=96$, $p=.005$). This means that memory aiding has a negative effect on the number of visits that subjects made. Without aiding the subjects begins with a relatively large number of visits and this discrepancy becomes even larger per session relative to starting with aiding. Removal of the aiding is accompanied with an increase of the total number of visits and introducing memory aiding is accompanied with a reduction in the total number of visits.

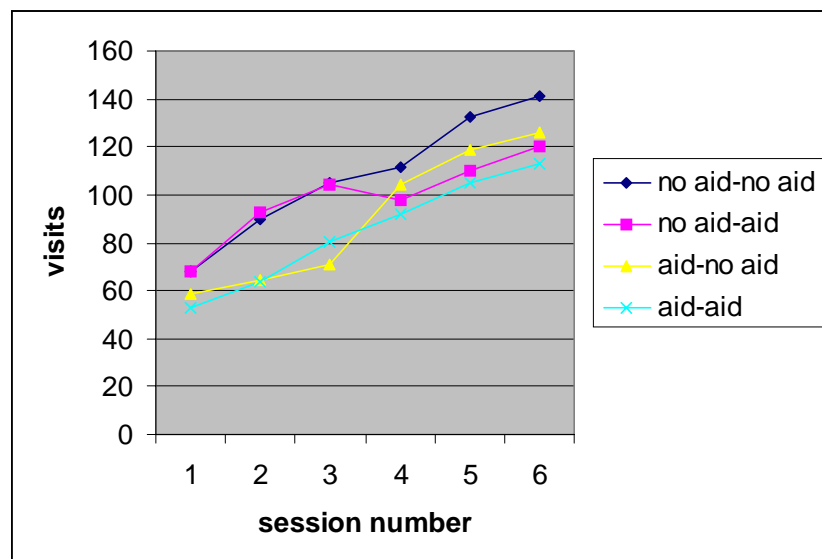


Figure 3.4: Overview of subjects decision speed score: Number of decision rounds (visits) per session for each of the experimental conditions

3.4. Amount of info. As mentioned earlier, subjects have the opportunity to search for information about the value of the prospects. Figure 3.5 and Table 3.2 (see appendix V) present the mean amount of purchased information per visit. The figure and table show that, in general, subjects search for only a limited amount of information, the average is always less than 1 request per visit. As can be seen in the figure subjects who start

in the no aid condition increasingly search for more information, up to twice as much as the subjects who started in the aid condition. Indeed, this interaction of the begin aid manipulation with practice is highly significant ($F=3,791$, $df=96$ $p=.004$). There is also a main effect of practice ($F=8.54$, $df=5$, $p<.001$) and a main effect of the begin aid manipulation ($F=.308$, $df=96$, $p=.907$).

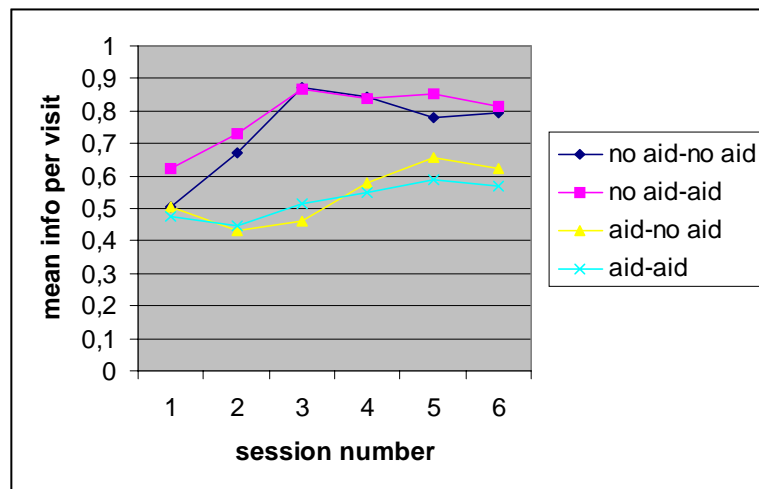


Figure 3.5: Overview of subject's info uptake: Mean amount of info per visit for each of the experimental conditions

($F=8.50$, $df=1$, $p<.001$) while there is no practice interaction with the end aid manipulation and nor a main effect of the end aid manipulation ($F=.917$, $df=5$, $p=.473$). This means that there is a strong transfer effect of the begin aid condition on the amount of information that is purchased in the last 3 sessions: Starting in the aid condition subject search for relatively little information in the first three sessions, this does not change when aiding is removed in the last three sessions. On the other hand, starting in the no aid situation subjects search for more information in the first three sessions, however this does not change when aiding is introduced in the last three sessions.

3.5. Type of information. A final analysis concerns the pattern of information types that subjects purchase. As mentioned earlier subjects can search for information concerning the value of local or global alternative prospects (which we label: 'local' vs 'global'). They can also purchase information concerning the future value of the prospect (which we will label: 'future'). In contrast they can focus only on the value of the prospect which is currently being offered (labeled: 'imminent'). Alternatively, subjects can purchase the information only after the decision is already made (labeled: 'after'). A general overview of these results is presented in table 3.4 for each of the information types over all sessions and groups. In the table figure it can be seen that subjects generally neglect information concerning the future value of prospects.

Table 3.4: Mean and median (in brackets) amount of purchased information type

Info type	Mean
After	.36 (.29)
Imminent	.17 (0)
Future	.01 (0)
Local	.03 (0)
Global	.05 (0)

Subjects generally do not search for information concerning local and global alternatives. These three types of info are hardly ever purchased; the means are .05 or less while the median value for each of these three information types is zero. There is too little variation in the data for these variables to examine to differences over sessions and groups. When subjects do buy info it is about the imminent option (mean value .17), but more particularly about the value of the products they have just purchased. That is: after the decision is made (mean value .36). Indeed, subjects buy more information after the decision than before the decision ($t=6.13$, $df=103$, $p<.001$).

4 Discussion

Several conclusions can be drawn from this experiment on practice and memory aiding in C2 decision processes. First, practice has a positive effect on decision performance. It increases the success score for each decision. It increases the overall decision speed and decreases the overall failure score. Second, memory aiding does not influence the mean success score for each decision, nor does it influence the overall failure score. If any: aid spoils performance in the subsequent non-aid session (negative transfer). Moreover, it slows down the decision speed considerably. Third, practice leads to an increase of the amount of information that subjects search for. Fourth, memory aiding in general has a negative effect on the search for information. Fifth, the amount of information that subjects searched for per decision is less than 1 request. Sixth, subjects search more information after the decision than before making a decision. Seventh, subjects do not look for alternatives and neglect information about future developments of the system.

One of the earlier mentioned problems of the experimental studies with simulated C2 environments was that the C2 task should neither be too complex nor too simple. From subjects' performance measures of success, failure and speed it can be concluded that in general the C2 task used in this study was not too easy nor too difficult. The failure scores (bankruptcies) dropped considerably over sessions, but subjects failed until the last session. No effects of aiding were found on the failure scores. In each session subjects' decisions were more successful in terms of the success score (profit) and this levels off, but not until the final session, which suggests that they reached their success 'peak' only in the last session. As mentioned earlier the mean success score per visit (profit) was used in the statistical analysis. A suggestion for a further more detailed analysis would be to use a time series analysis on the profit score. This may provide further insight into the more detailed differences between the experimental groups.

However, one should be careful in concluding that aiding has no effect on overall success performance, since aiding in general reduced the total number of decisions that were made in a session. This means, that if a subject makes 2000 profit units per decision and his total number of decisions decreases, then that subject will end up with less profit. Thus, aiding may not seem to influence the mean success score per decision, it does have however a negative effect on the overall success score per session.

The effect of aiding on decision speed, i.e. the total number of decisions made in a period of time is both intricate and clear. In Figure 3.3 it can be seen that in general subjects were able to make more decisions per session up to the very last round. As a result of aiding subjects make less decisions relative to the no-aid group, about 20-25% less decisions per session. Introducing aiding to a group that was used to work without aiding resulted in a decline in the number of decisions. Similarly, removing the aid from a group that was used to work with it resulted in a clear increase in the number of decisions made. There did not appear differences in the mean success score per decision. It is plausible that memory aiding may lead subjects to spending too much time in trying to structure the decision problem. They may spend relatively much time in the uptake of information and the constructing of problem representation. Subjects may also grind too much on the information and actually think too much, at least too much relative to what the dynamics of the situation allow. If that is indeed the case, memory aiding turns out to be more like a burden than a support.

This points out a warning for introducing aiding techniques in decision making environments in general without sufficient research. An obvious means of help may become a burden in disguise. Many subjects in the no aiding condition mentioned: 'I can never do this task without paper and pencil'. Although this is only a qualitative observation, it illustrates our point. It may sound trivial that supporting working memory, by giving subjects some more storage capacity should be helpful; yet the results show that it was otherwise and

deteriorated human decision performance. Therefore, potential decision support techniques should first be tested in simulated lab studies before they can be applied in active service. It could be that providing more storage capacity promotes more extended considerations which are actually not helpful.

An important question in this study concerns the amount of information subjects require to know about the value of a certain prospect. In the simulation the prospects were buy offers at a trading centre at a particular moment. The subjects were left in continuous uncertainty about the local, global and future values of the prospects. In order to remain updated constant information has to be obtained on changes in supply and demand. In our C2 task simulation subjects evaluated three prospects per decision round. Our results show that subjects purchased on the average less than 1 piece of information concerning the local, global or future values of all three prospects. Although the results show that practice does increase and almost doubles the total profit, the conclusion must be that on many occasions the subjects make a decision without searching for any value of prospects at all.

Memory aiding had a negative effect on the amount of information that is searched for. More specifically, as Figure 3.4 shows, it is what they receive first, aiding or no aiding, what matters most. Subjects who started with aiding, searched for relatively less information compared to subjects who started without aiding. Removing aiding, however, did not increase the amount of information search. And vice versa, introducing memory aiding did not reduce it. How can this pattern of results be explained? First, it is plausible to relate these observations to the hypothesis that memory aiding makes the decision maker spend too much time in problem structuring. Following this conjecture, it is likely that subjects who start with aiding spend too much time or think too much about each piece of information they purchase. As a consequence they have less time to purchase information and consequently the mean amount of purchased information is less compared to the no aid condition. Second, once the subjects have adopted a certain information search profile, it is plausible to assume that adding or removing the memory aiding does not lead to a structural change in information search strategy. What happens is that both groups simply continue using their adopted search strategy, but become either slower or faster. In the condition of removal of the aid subjects will start making more decisions, and in the condition of adding the aid the number of decisions decreases.

The results on the information search profiles show a final important finding. Table 3.5 shows that subjects do not search for local or global alternatives; nor do they search for information concerning the future value of a prospect. It is striking that subjects neither searched for alternative options nor seemed to incorporate information about the future value of prospect into their decision making process. Such information could have given further insight on potential alternatives or future disasters or fortunes, but it was practically not searched for. Thus one could argue that subjects did actually not decide, since decision making requires that subjects choose between alternatives (Baron, 1994; Hogarth, 1987). In contrast subjects mainly evaluated available opportunities. They decided by evaluating the prospects one at a time, and rejected or accepted each option sequentially. Moreover, they bought more information after than before the decision. In other words: the general decision strategy can be described as: satisfy and justify. With a minimal amount of information they choose the 'first acceptable' option and used information afterwards to justify the decision. Such a strategy may lead to sub-optimal results: indeed most subjects had a significant failure score in the first sessions. Due to practice, however, subjects successfully adapted their strategy to the dynamics of the environment.

An important question is why subjects downgrade information about the future developments of the system. Why is imminent information overvalued, or is it really? One explanation is that the task was too complex. To incorporate the future values of prospects may have been just beyond the processing capacity of a single individual, especially under time stress. Another explanation is that subjects fail too see the importance of future developments. A third and not unrelated explanation can be found by relating this finding to the literature on time preference and inter-temporal choice (Loewenstein & Elster, 1992; Roelofsma, 1996). A well described phenomenon in this literature is that subjects downgrade future consequences. Immediate outcomes are overvalued because subjects view the future as uncertain. The conjecture has not been extended to the evaluation of timing of information, or the resolution of uncertainty of information. It is plausible, that subjects do not search for future information because they perceive this information to be less certain.

Our results provide some hints with regard to some major problems that could be faced by operators of complex systems, like military C2. In this respect it offers some guidelines to those issues that should be focussed on and that require further study. A major result of this experiment is the finding that memory aiding may slow down the decision making process and have a negative effect on the overall decision performance. Aiding techniques should therefore be carefully and systematically examined in experiments with simulated C2 task environments before put to practice. Another major finding of our experiment was that subjects attend only to a limited number of information concerning the value of a prospect. Although practice was helpful, the improvement was not striking since most of the information was used to justify the decision. As a result of technological advances there is a tendency to feed C2 systems with increasingly more information. Giving our findings it is more reasonable to carefully examine each piece of information that is put in the system. It is important to know whether and how the operator uses this information in the decision process before it is fed into the system. The presence of abundant information may deteriorate decision performance and the information search in a number of ways. The effects of redundant and irrelevant information in C2 as well as the above mentioned effects of the time frame of the information (imminent vs remote) definitely requires additional research that could not be conducted in the context of the present project.

References

- Andriole, S.J. and Hopple, G.W. (1982). They're only human: Decision makers in command and control. Signal, march, 1982.
- Baron, J. (1994). *Thinking and Deciding*. Cambridge University Press, New York.
- Broadbent, D., Berry, D. & Gardner, P. (1990). Experiments on the role of action and of advice. Report 2PR1DEB from the Bonn group of the KAUDYTE project 'Knowledge acquisition and use in complex and dynamic task environments'.
- Brehmer, B. (1987). System design and the psychology of complex systems. In: J. Rasmussen and P. Zunde(Eds.), *Empirical foundations of information and systems science*. New York: Plenum.
- Doorne, H. van (1986). *MILSIM: a computer simulated game for C² systems*. Internal report TNO-IZF.
- Dörner, D. (1987). On the difficulties people have in dealing with complexity. In J. Rasmussen, K. Duncan and J. Leplat (Eds.), *New technology and human error*. Chichester: Wiley.
- Ebbesen, E. and Konecni, V. (1975). Decision making and information integration in the courts: The setting of the bail. *Journal of Personality and Social Psychology*, 32, 805-821.
- Einhorn, H. (1974). Expert Judgment: Some necessary conditions and an example. *Journal of Applied Social Psychology*, 59, 562-571.
- Festinger, L. (1957). *A theory of Cognitive Dissonance*. Evanston, Ill. Row, Peterson.
- Flin, R., Salas, E., Strub, M. & Martin, L. (1997). *Decision Making under Stress*. Emerging themes And applications. Ashgate, Aldershot.
- Geath, G.J. and Shanteau, J. (1984). Reducing the influence of irrelevant information on experienced decision makers. *Organizational Behavior and Human Decision Processes*, 33, 263-282.
- Gigerenzer, G., Hell, W. and Blank, H. (1988). Presentation and content: The use of base rates as A continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 5, 13, 25.
- Gilovich, T. & Medvec, V.H. (1995). The experience of regret: What, when and why? *Psychological review*, 101, 379-395.
- Hammond, K.R., Hamm, R.M., Grassia, J. & Pearson, T. (1987). Direct Comparison of the efficacy of intuitive and analytical cognition in Expert Judgement. *IEEE Transactions on systems, man and cybernetics*, Vol, SMC-17, no 5, 753-770.
- Hogarth, R.M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. *Psychological Bulletin*, 90, 197-217.

- Hogarth, R.M. (1987). *Judgement and choice*. Wiley and Sons. Chichester, New York.
- Huber, O., Debeutz, A., Pratscher, J. And Quehenberger, I. (1990). Perceived control in a multistage decision task. *Journal of Behavioral Decision Making*, 3, 123-36.
- Kerstholt, J. (1996). *Dynamic decision making*. Dissertation, TNO-IZF.
- Lusted, L.B. (1977). A study of the efficacy of diagnostic radiologic procedures: Final report on diagnostic efficacy. Chicago: Efficacy study committee of the American College of Radiology.
- Loewenstein, G. & Elster, J. (1992) *Choice over time*. Russell Sage Foundation.
- Kleinmuntz, D.N. (1985). Cognitive heuristics and feedback in a dynamic decision environment. *Management Science*, 31, 680-702.
- Mynatt, C.R., Doherty, M.E. and Tweney, R.T. (1977). Confirmation Bias in a Simulated Research Environment: An Experimental Study of Scientific Evidence. *Quarterly Journal of Experimental Psychology*, 29, 85-95.
- Roelofsma, P.H.M.P. (1996) Anomalies in Intertemporal Choice, *Acta Psychologica*, 93, nos 1-3, 5-23.
- Simon, H.A. (1957). *Models of man: Social and rational*. New York: Wiley.
- Svenson, O. & Benthorn, L.J. (1992). Consolidation processes in decision making: Post-decision changes in attractiveness of alternatives. *Journal of Economic Psychology*, 13, 315-327.
- Tversky, A. & Kahneman, D. (1982). *Judgement under uncertainty: Heuristics and biases*. Plenum Univ.Press.
- Wickens, C.D. (1992). *Engineering psychology and human performance* (2nd edition). New York: Harper Collins.
- Zsombok, C.E. & Klein, G. (1997). *Naturalistic Decision Making*. Lawrence Erlbaum Associates, New Jersey.

APPENDIX I

In this experiment you will participate in a management command control simulation. There are six different commodities that can be traded and six different trading centres that can be visited. These centres are located at different stars centres in space. The commodities are: Food , computers, deuterium, platinum, machines and preciosa. Each centre deals with all six goods, but the prices are different at each centre and they are continuously changing as well. Moreover, a centre either buys or sell each commodity, but not both sell and buy the same product.

You are the commander of a fleet that transports commodities from one centre to the other in a competitive environment with another fleet. You will buy commodities in centre that supplies them and sell commodities to centres that have a demand. Your ultimate goal is to make as much profit as possible. For this reason you have to buy at the lowest price and sell at the highest price.

As mentioned you are the commander of the fleet. You can use the ships to transport the goods from one centre to the another. When a ship arrives at a centre you can first sell the commodities that are demanded. Then, you can buy commodities that are offered at the centres and that you want to transport to another centre. When you have finished your transactions you have to decide your next travel destination, that is you have to indicate which centre you will visit next. However, you can not leave a centre unless you have sufficient fuel supplies. So, pay attention to having sufficient capital to afford the travel costs. These travel cost are influenced by:

- a) The fuel price (this may change over time)
- b) The weight of the ship's cargo. The transaction are made in a unity called 'kugel'. However kugels differ in weight for each commodity. You can see the current ship's weight of the commodity on the computer screen.
- c) The travel destination. The longer the distance the higher the travel costs will be. You can look at your star centre map for the distances between centre. The current travel costs per 100 star miles is presented on your command and control screen.

You will control your fleet with a computer communication network that is directed by your command and control screen. The negotiations with the centres will also be with the command and control screen. Both sell- and buy prices are continuously fluctuating. The quantities expressed in kugels vary over time as well. The indicated prices are dependent on the demand on a particular moment. When a centre has bought an amount the demand of that particular product will decrease and the prices will also be decrease. That is, both prices and quantities fluctuate depending on the transaction pattern.

To simplify your decisions you can purchase information about quantities and prices. You can also ask information about the development of prices and quantities, supplies and supplies and contextual information of the centres by the news bulletin.

There are three types of information:

- a) *Centre information.*

By providing the name of a star centre you will receive on your second command and control screen an overview of all the current prices and quantities that are being traded in that centre.

- b) *Commodity information.*

By providing the name of a commodity you will receive an overview of current prices and quantities of a commodity at all centres.

- c) *Bulletin information.*

Here you can receive information about 1) future developments of prices and quantities. The system is dynamic and prices and quantities change continuously. In some periods of time fluctuation in prices is large, at other times it is small. 2) Information about the development of the fuel price which can increase

or decrease. The travel costs are influenced by the fuel price. 3) Information about the future costs of information. 4) Information about the travel speed of the ship. The travel speed of your ship can be influenced by the condition of the machines and the environmental conditions. This info can help you decide whether or not to travel for long or short distances.

It is important to realize that the travel costs can increase considerably. Again, they are influenced by the weigh of your cargo and your travel destination. Their must be sufficient capital on board of a ship to pay these costs. If the travel costs exceed the cash that is available on a ship there will be two options for the commander:

- 1) change your travel destination
- 2) Resell products to the current starcentre. But please pay attention to the fact that the products are being resold only for half the original price.

It is important that to realize that you keep on trying to make as much profit as possible. There is a competing team which acts as an enemy of your company!

APPENDIX II

Table 2.1: Development of prizes, quantities, and game model

Development of prize, demand and supply:

Concerns.. Transaction	Prize on centre	Prize on other centre	Demand on Centrum	Demand on other centra	Supply on centrum	Supply on other centra
Centrum sold	Up	Down	n.a.	n.a.	n.a.	Up
Centrum bought	Down	Up	Down	Up	n.a.	n.a.

Some important formulae of the model behind the game:

Centrum buys/sells:	$0.6 * \text{stock}$
Development of stock:	$Q_n = Q_o - \text{PARQ} * N * \text{RAND}$
Development of prize in centre:	$P_n = P_o + \text{PARAM} * N * \text{RAND}$ (Buy: $\text{PARAM} = -\text{PARP}$) (Sell: $\text{PARAM} = \text{PARP}$)
Development of prize in the two other buying or selling centre:	$dP = \text{abs}((P_o - P_n)/2) * \text{RAND}$
Development of stock in the two other buying or selling centre:	$dQ = \text{abs}((Q_o - Q_n)/2) * \text{RAND}$
Travelling-time of the ship:	$\text{weight} * \text{PART} + \text{distance} * 0.1$
Travelling-costs:	$\text{weight} * \text{distance} * \text{prize of gas} + 1000$
Prize of information:	PARI
Prize of fuel (gas):	PARG
Randomising:	Rand = random selection of (0.6, 0.8, 1.0, 1.2, 1.4)

During the simulation the parameters PARP, PARQ, PARG, PART en PARI are fluctuated according to the parameter table.

N = merchandized number n = new value o = old value

Each centre sells 3 commodities and buys 3 commodities.

Type of commodity to sell or buy differs on each centre

APPENDIX III

Table 3.1: Subjects success score: Mean amount of profit per sessions for the experimental groups and sessions. The medians are printed in brackets

SESSION	1	2	3	4	5	6	TOTAL MEAN
No aid-No aid	-1290.63 (-1594.51)	137.84 (380.82)	1102.90 (1092.95)	1242.77 (1212.27)	1818.55 (2065.88)	2175.32 (2078.62)	864.46
No aid- Aid	-753.82 (-1431.33)	-129.01 (-705.22)	770.30 (1086.44)	1137.33 (1345.25)	1906.18 (1868.75)	1737.59 (1677.58)	778.10
Aid-No aid	-1197.97 (-1331.62)	-132.50 (-711.30)	482.91 (-24.41)	705.10 (30.80)	1162.31 (1411.33)	1038.53 (1518.55)	343.06
Aid-Aid	-1095.65 (-1499.74)	-68.23 (-816.05)	258.60 (326.39)	878.53 (551.42)	1374.34 (1608.79)	1840.71 (1739.91)	531.38
TOTAL MEAN	-1084.52	-47.98	653.68	990.93	1565.35	1698.04	

APPENDIX IV

Table 3.2: Subjects decision speed score. Mean amount of decision rounds (visits) per sessions for the experimental groups and sessions. The medians are printed in brackets

SESSION	1	2	3	4	5	6	TOTAL MEAN
No aid-No aid	67.73 (62.50)	89.46 (89.50)	104.65 (98.50)	111.42 (102.50)	132.50 (126.50)	141.04 (131.50)	107.8
No aid- Aid	68.08 (62.50)	92.73 (99.50)	104.62 (88.50)	97.46 (91.00)	110.27 (104.50)	119.92 (110.00)	98.85
Aid-No aid	58.46 (50.00)	64.08 (64.00)	70.92 (67.50)	104.19 (95.00)	118.65 (117.50)	125.77 (123.00)	90.35
Aid-Aid	52.73 (46.00)	64.00 (60.00)	80.19 (73.00)	92.19 (84.50)	104.65 (93.00)	113.04 (103.50)	84.47
TOTAL MEAN	61.75	77.57	90.10	101.32	116.52	124.94	

APPENDIX V

Table 3.3: Mean amount of information that subjects search for about the value of a prospect for the experimental groups and sessions. The medians are printed in brackets

SESSION	1	2	3	4	5	6	TOTAL MEAN
No aid-No aid	.51 (.53)	.67 (.67)	.87 (.74)	.84 (.82)	.78 (.67)	.80 (.74)	.75
No aid- Aid	.62 (.50)	.73 (.66)	.87 (.77)	.84 (.78)	.85 (.72)	.81 (.69)	.79
Aid-No aid	.51 (.49)	.43 (.22)	.46 (.20)	.58 (.51)	.66 (.59)	.62 (.63)	.54
Aid-Aid	.48 (.40)	.45 (.40)	.51 (.43)	.55 (.48)	.59 (.50)	.57 (.44)	.53
TOTAL MEAN	.53	.57	.68	.70	.72	.70	

The Risk of Human Error: Data Collection, Collation, and Quantification*

J W Chappelow

Centre for Human Sciences
DERA Farnborough
Farnborough, Hants. GU14 0LX
United Kingdom

Summary: Human performance poses significant problems in system reliability assessment. Are realistic assessments of safety in systems involving humans possible? Can human performance be quantified? What aspects of human performance are predictable? Practical experience in the field of aviation safety suggests some answers to these questions.

Introduction: This is a historical account of a variety of projects concerned with human error in aviation. As a summary of personal experience it is necessarily partial, in both senses; that is to say it is an incomplete and biased view of human reliability. It may, nevertheless, cast some light on the themes of the workshop: Can the safety implications of human performance be addressed rigorously? What should be predicted? Is meaningful quantification possible?

Classification 1: Psychologists have assisted Royal Air Force Boards of Inquiry since 1972. By 1982, enough reports on aircraft accidents had been collected to allow a first attempt at organising the data and seeking patterns. The classification scheme devised then had no particular theoretical bias, was simply organised, and allowed the most prevalent contributory factors to be identified.^{1,2,3} They are shown in Table 1 grouped under arbitrary headings.

On the basis of this analysis, research projects addressing personality issues and cognitive failure were undertaken.⁴ Although some interesting findings resulted, neither project led to practical innovations to reduce risk beyond general guidance given to flying supervisors in flight safety courses. It is interesting to note, in retrospect, that both projects addressed individual susceptibility to particular types of error. This was probably a reflection of political rather than technical realities at the time. Although the role of design and organisational factors in human error was well recognised, there was still a remnant of “blame culture” to be overcome.

Table 1: The most common contributory factors

Aircrew		System	
Inexperience	23%	Training & briefing	25%
Personality	21%	Administration	23%
Life stress	14%	Ergonomics	22%
Social factors	11%	High workload	14%
Immediate causes			
Acute stress	26%	Inappropriate model	16%
Distraction	20%	Visual illusion	10%
Cognitive failure	17%		

* © Defence Evaluation and Research Agency

Two sorts of insight resulted from these initial efforts: Identification of the more important contributory factors; and the recognition that both the size of contribution to overall risk and the tractability of the problem were important in determining where to invest remedial effort. Tractability and quantifiability turned out, initially at least, to be associated.

Quantification 1: Few emergencies in aviation require an immediate response. Helicopters have more than their fair share of those that do. A prime example is total power failure. It requires an immediate reduction in collective pitch. How long the pilot has to achieve this depends on the inertia in the rotor disc, and this is an issue of relevance to the certification requirements for helicopters.

Reaction times are relatively easily and objectively measured. They have long been a mainstay of experimental psychology. Unfortunately, it is difficult to generalise with convincing precision from laboratory studies, however sophisticated, to real world situations. It was necessary to resort to flight simulator experiments. Figure 1 shows some of the results for three helicopter types: means and 90th percentiles for detection time (the interval between the emergency onset and the first indication of an appropriate response) and response time (the time taken to complete the action).⁵ It seems that reaction times even for well-practised responses to easily identified conditions can be surprisingly long, particularly when the normal variability of behaviour is taken into account.

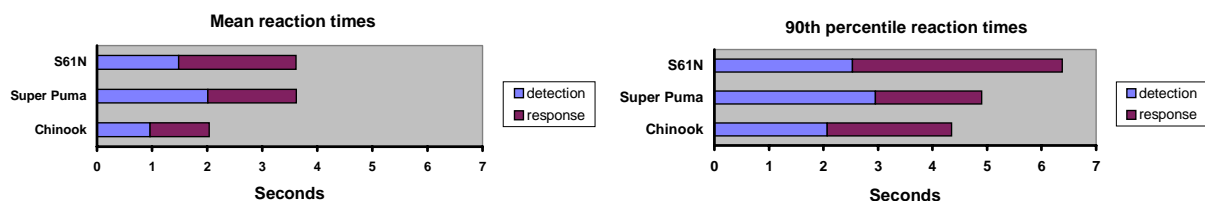


Figure 1: Reaction times to total power failure

These results have a direct bearing on the mechanical design of helicopters. When the probability of total power failure is known, they allow the risk of an unfavourable outcome to be estimated in a way that allows cost-benefit analysis to inform design decisions.

A variety of helicopter emergencies were addressed in this study.⁶ Although some instructive differences were found, similar results were obtained in several cases and in a dissimilar case – an untrained-for and (from the designer’s perspective) unpredictable control malfunction in a fixed-wing aircraft. The findings do provide general guidance on the reaction times to be expected in a range of situations within aviation, at least. It is also clear that there are limits to this generalisability, and it is not clear how wide a range of similar studies would be required to provide comprehensive guidance on reaction times in real situations. Such guidance would, however, be valuable to system designers and regulators, and could be relatively easily obtained. A sensible first step would be the classification of situations in terms of the types of task and responses involved.

Quantification 2: The UK Low Flying System (UKLFS) is uncontrolled airspace from ground level to 2000ft. It is used by a variety of civilian aircraft – hang-gliders, microlights, gliders, fixed- and rotary-wing light aircraft – as well as military helicopters, transports, and fast jets operating at speeds in excess of 400kt. All operate on the “see-and-avoid” principle. The risk of random mid-air collision is real. Collisions involving two fast jet aircraft not surprisingly provide the most numerous examples of this risk. They also represent an extreme and, therefore, relatively simple case, the most important features of which (the psychophysical aspects) can be modelled sufficiently precisely to allow useful predictions to be made.

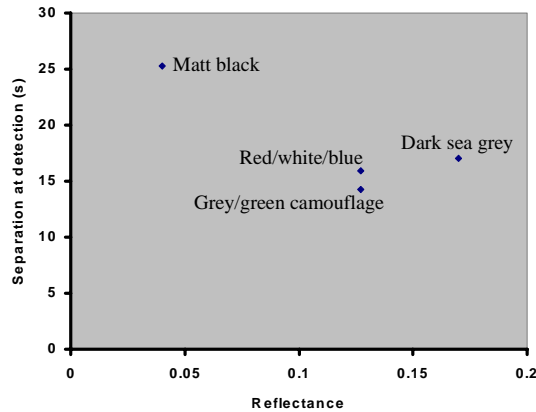


Figure 2: Flight trial results (paint schemes)

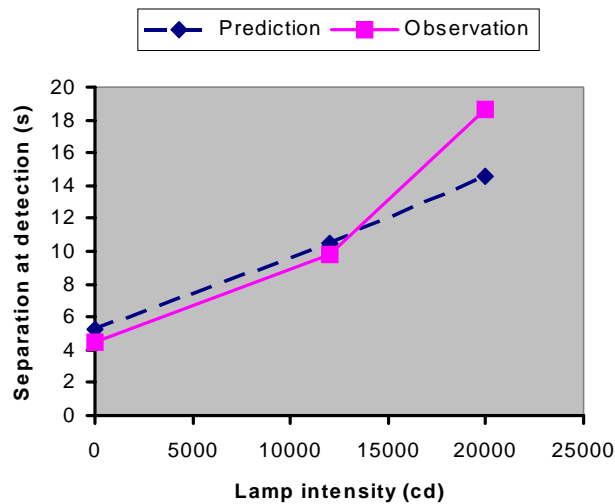


Figure 3: Flight trial results (lamps)

An initial, approximate attempt at such modelling suggested advantages for black paint schemes and for very bright, fixed, steady lights – as opposed to the high intensity strobe lights commonly fitted to aircraft.⁷ It also allowed the risk reduction achievable through electronic collision warning systems to be estimated. Flight trials confirmed the predictions (Figures 2 and 3 show sample results), and supported refinement of the model.^{8,9,10}

In a further project, the psychophysical model was combined with a computer simulation of activity in the UKLFS.¹¹ It was necessary to collect a large amount of data to support this modelling exercise (Figure 4). The resulting predictions were validated against reported conflict rates (from the Joint Airprox Working Group) and the historical record of collisions. The principal predictions (one fast jet–fast jet collision every two years and one military–civilian collision every six years) have continued to prove tragically accurate. However, the estimates of the effectiveness of remedies such as paint schemes and collision warning systems derived from the model have informed the continuing debate on safety in the UKLFS and influenced policy decisions.

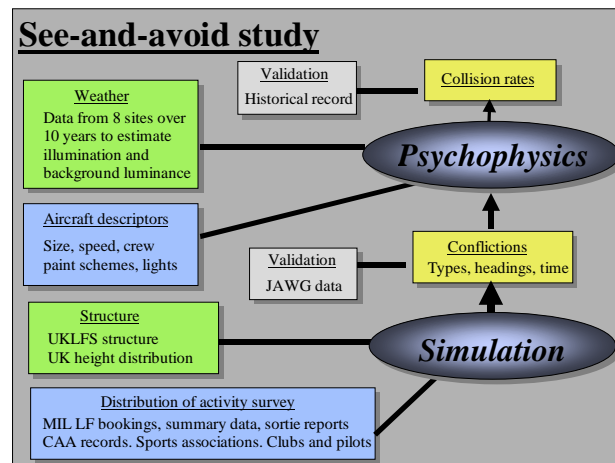


Figure 4: Construction of a predictive model

Classification 2: The need for a precise and useful classification scheme for human error and its underlying causal factors has become more pressing. Involvement with NATO RSG 25 allowed a less aviation-specific model to be drafted, and this formed the basis of a recent project aimed at developing a causal factors database for both the human factors and the engineering domain within military aviation.¹²

Although computer systems are changing the picture, the engineering domain has been characterised by a plethora of subsystems and components each of which has a limited range of functions (usually only one each) and only a few ways of failing. The human factors domain is characterised by one component (*Homo sapiens*) which serves a multitude of goals (rather than simple functions), and has many ways of failing.

Accident and incident databases in aviation have tended to follow a model appropriate to the engineering domain, and have been relatively uninformative as to the causes of human error. Indeed, there is a parallel between the traditional engineering approach (identify the defective component and replace it) and the old-fashioned approach to human error (find out who is to blame and punish them). The new database allows for simple classification of human errors and a flexible, hierarchical coding of causal mechanisms designed to identify all types of contributory factors (Figure 5 is an outline). By imposing a similar model on the engineering domain (Figure 6), a different perspective on the causes of mechanical failure has been obtained, which has resulted in at least one unexpected insight concerning the detection of problems between flights.

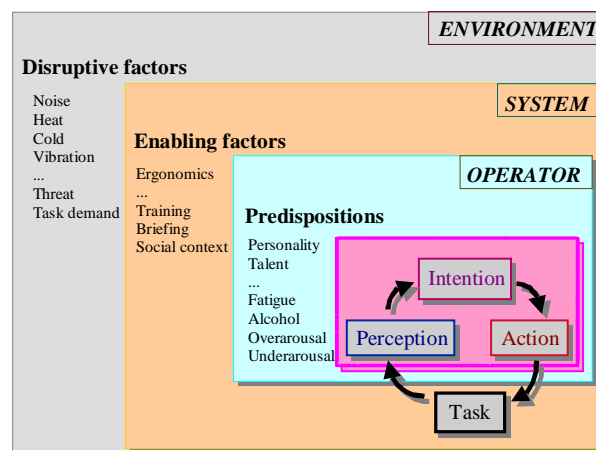


Figure 5: Outline human factors classification

The database has also been used to prototype a risk analysis system. By using historical data to estimate the quality of underlying causal factors and the strength of their influence on failure mechanisms, relatively objective sensitivity analysis has been made possible. Comparison of Figures 7 and 8 shows the broader

perspective and added complexity derived using this approach in comparison with a similar procedure based on experts' opinions when both approaches were used to analyse the factors underlying one type of accident.

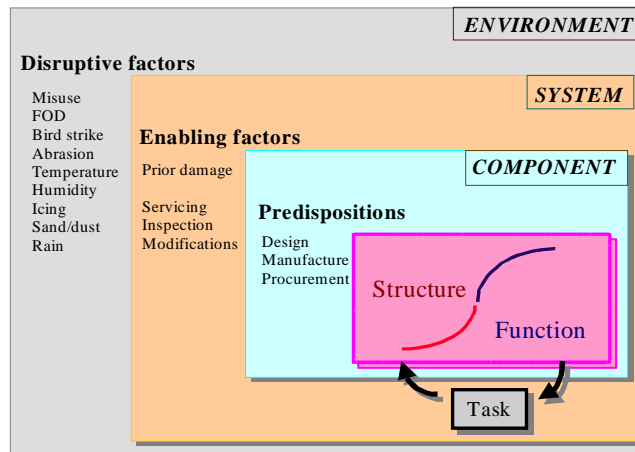


Figure 6: Outline engineering classification

Sensitivity analysis applied to the whole range of accidents has revealed the strongly influential character of social factors in military aircraft accidents – a fact not evident in simpler analyses. These factors can be addressed via training programmes – a relatively cheap and immediate option in comparison with other remedies for error such as hardware modification, for example. The fact that they are influential as well as relatively tractable makes them an important target in flight safety programmes.

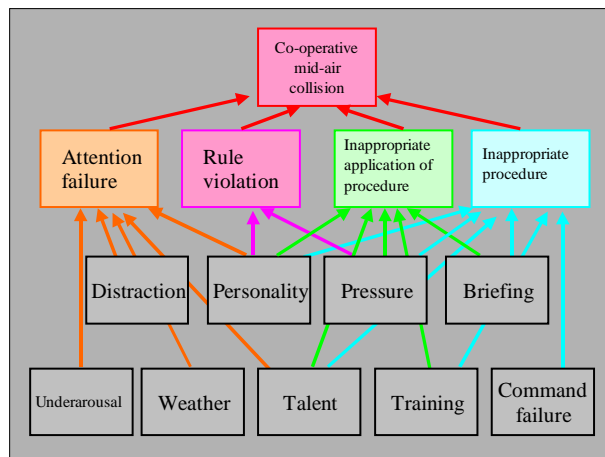


Figure 7: Influence diagram generated by experts

A recent review of social factors in accidents was intended to refine the RAF crew resource management training programme by identifying social factors influencing ground-based as well as airborne activity.¹³ The factors identified include not only communication problems and decision making biases already known to affect small teams, such as the “risky shift” phenomenon, but also organisationally-induced tendencies to more risky behaviour.¹⁴ There may be parallels here with the risk conservation behaviour reported in the road safety context.¹⁵ It is certainly clear that, whatever the intention behind the design of a system, individual operators, small groups or teams, and even whole organisations may use it for aims undreamed of by the designer. Individuals derive status, satisfaction, fun, even thrills from the use of systems, and teams and organisations may similarly add to or even subvert the formally defined purpose. The social contexts that promote these parallel or supplementary purposes deserve attention since they define a whole category of risk otherwise ignored.

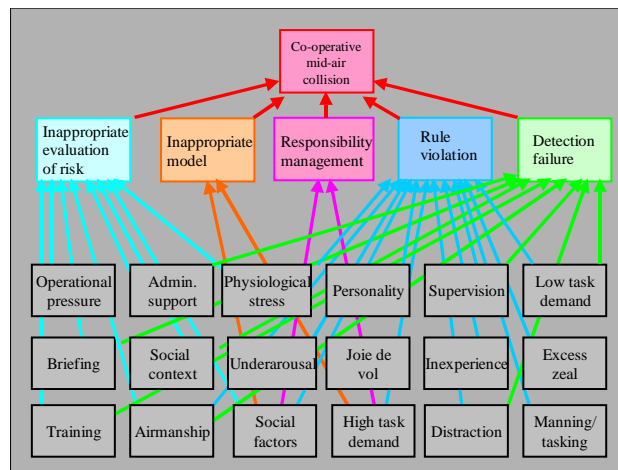


Figure 8: Influence diagram based on historical data

Incident data (and Quantification 3): Accident investigations provide a rich source of detailed information on risk and reliability, but accumulate only slowly. Incidents (near misses) are more numerous, and logically deserve equal attention since in principle, they could provide much more data. Confidential incident reporting schemes have been introduced in many industries besides aviation as a way of increasing the amount of data collected. Recent experience in the RAF suggests that open reporting may be even more effective in uncovering unsuspected problems. Such a system requires the prior establishment of an appropriate organisational culture so that a guarantee of immunity from punishment for honest mistakes will be accepted at face value. There always remains, however, a problem in assessing the magnitude of a reported risk, as the following example illustrates.

Ejection seats are intended to save life, but are potentially lethal. Most are made safe by inserting mechanical barriers into the firing mechanism - usually pins. In thirty years the RAF has recorded two fatal accidents involving ejection seat pins. In one case the seat was safe when it should have been live (a Type 1 error). In the other case it was live when it should have been safe (a Type 2 error). Only eleven *incidents* involving seat pins were formally reported in the same period.

Shortly after the introduction of open reporting, a change in the procedures used at one flying station resulted in several Type 2 errors, which were reported. As interest focussed on this particular location, a small number of Type 1 errors appeared as well. These could not have been caused by the change in procedure. They appeared to have come to light simply because of the locally heightened interest in seat pins procedures. On this basis it was suggested that other aircraft types and flying stations must also be experiencing seat pins errors. This encouragement produced a small crop of reports of both types of error. At this stage, it was clear that a problem of unknown magnitude had been uncovered. To estimate its prevalence, questionnaires were used to capture all seat pins errors occurring during one month.

The results of this survey suggest that about 100 Type 1 errors and 200 Type 2 errors are made every year in the RAF. These potentially lethal errors have presumably been occurring since the introduction of ejection seats, and have barely come to official notice except when accidents occurred. To obtain a realistic estimate of the error rate, it was necessary not only to advance beyond mandatory and confidential incident reporting programmes, but also to collect data on this specific topic for a defined period.

A simple count of the frequency of an error is not enough to gauge its importance. Combining the probability of a Type 1 error with the probability (obtained from accident data) that ejection will be required enables the risk of a fatal outcome to be calculated. This gives real meaning to conventional reliability standards such as 1 fatality in 10^6 or 10^7 sorties. On present estimates, the risk due to Type 1 errors warrants serious consideration of modifications to current operating practices and a re-evaluation of the general approach in future ejection seat designs. If a different standard were adopted, 1 in 10^9 for example, the implications would be far more severe, and immediate, drastic action would be required.

Conclusions: The practical experience described here suggests some conclusions that might possibly have general relevance. Meaningful quantification of human performance in a form that is useable in reliability assessment does seem to be possible. It is, however, probably significant that the two major examples given involve relatively simple aspects of behaviour – reaction times and visual psychophysics. In both examples, the stimulus conditions and the required responses were closely defined. The third (ejection seat) example also involves relatively simple behaviour. In tasks demanding more interpretation or complex decision making, the challenge of meaningful and rigorous quantification may be considerably more daunting.

Although laboratory studies can provide a rigorous understanding of specific error mechanisms, a realistic appreciation of the potential for human error can only be obtained by close scrutiny of real systems. This implies thorough investigation of the human factors aspects of accidents and the collection of data on “near-miss” incidents. Such data are, of course, useless unless organised and collated in a way that illuminates failure mechanisms and allows practical remedies to be devised. We have demonstrated that classification can be developed to the point of permitting relatively objective risk assessment. However, the ejection seat example demonstrates that considerable, focussed effort is required to obtain reliable estimates of error rates in the real world, and that reliance on accident statistics or conventional incident data alone is likely to result in a substantial underestimate.

Finally, although it is possible to quantify the probability of error in, say, dial reading or switch operation in a way that parallels reliability assessment of engineering components, this ignores important facts about human operators. They have goals rather than functions. Some of their goals are not those envisaged by system designers. Some are determined by characteristics of the teams they work in or of the organisation as a whole. These factors are also amenable to systematic analysis, possibly even to quantification. In addressing system reliability, we need to consider not just the artefact-system, or the man-machine system, but the whole system-complex.

References:

1. Chappelow, J. W. (1984). *Human error in aircraft accidents: A review of psychologists' reports on RAF accidents 1972-1982*. IAM Report 663. Farnborough, UK: RAF Institute of Aviation Medicine
2. Chappelow, J. W. (1988). *Causes of aircrew error in the Royal Air Force*. In *Human Behaviour in High Stress Situations in Aerospace Operations*. AGARD Conference Proceedings 458. Neuilly sur Seine, France: NATO Advisory Group on Aerospace Research and Development.
3. Chappelow, J. W. (1989). *Remedies for aircrew error*. IAM Report 664. Farnborough, UK: RAF Institute of Aviation Medicine.
4. Smith, A., Chappelow, J. W., and Belyavin, A. J. (1995). *Cognitive failures, focused attention, and categoric search*. Applied Cognitive Psychology, Vol. 9, S115-S126.
5. Smith, P. R. and Chappelow, J. W. (1995). *Pilot Intervention Times*. In *Human Factors in Aviation Operations, Proceedings of the 21st Conference of the European Association for Aviation Psychology (EAAP)* Vol 3 ed. R. Fuller, N. Johnson and N. McDonald. Aldershot, UK: Avebury Aviation.
6. Chappelow, J. W. and Smith, P. R. (1997). *Pilot intervention times in helicopter emergencies: Final report*. PLSD/CHS/HS3/CR97020/1.0. Farnborough, UK: DRA Centre for Human Sciences.
7. Chappelow, J. W. and Belyavin, A. J. (1991). *Random mid-air collisions in the low flying system*. IAM Report 702. Farnborough, UK: RAF Institute of Aviation Medicine.
8. Chappelow, J. W. and Belyavin, A. J. (1992). *A trial to assess aids to conspicuity*. IAM Report 723. Farnborough, UK: RAF Institute of Aviation Medicine.
9. Chappelow, J. W., Belyavin, A. J., and O'Connor, E. M. (1992). *A further evaluation of fixed steady lights as aids to conspicuity*. IAM Report 732. Farnborough, UK: RAF Institute of Aviation Medicine.
10. Chappelow, J. W., Belyavin, A. J., and Smith, P. R. (1993). *Aircraft conspicuity and paint schemes: A further trial*. IAM Report 747. Farnborough, UK: RAF Institute of Aviation Medicine.

11. Chappelow, J. W., Belyavin, A. J., Ruskell, L., Procopides, M., Smith, P. R., and O'Connor, E. M. (1997). *See-and-avoid operational analysis study*. PLSD/CHS/HS3/CR97029/1.0 DRA Farnborough, UK: DRA Centre for Human Sciences.
12. Chappelow, J. W., O'Connor, E. M., Johnson, C., and Blair, R. C. (1999). *Causal factors in military aircraft accidents*. DRA/CHS/MID/CR990245/1.0. Farnborough, UK: DERA Centre for Human Sciences.
13. Chappelow, J. W. and O'Connor, E. M. (1998). *Crew resource management training and military flight safety*. DERA/CHS/MID/CR980268/1.0. Farnborough, UK: DERA Centre for Human Sciences.
14. Snizek, J. A. and Henry, R. A (1989). Cited in Sutherland, S. (1992). *Irrationality*. London: Constable and Company Ltd..
15. Janssen, W. (1994). *Seat-belt wearing and driving behaviour: an instrumented-vehicle study*. Accident Analysis and Prevention, Vol. 26, pp 249-261. Pergamon, London.

Safety Culture – Theory and Practice

Patrick Hudson

Centre for Safety Science
Universiteit Leiden
P.O.Box 9555
2300 RB Leiden
The Netherlands
Hudson@fsw.LeidenUniv.nl

Abstract

Safety Culture is seen as a way of ensuring high levels of safety performance in organisations, in contrast to the systematic engineered management of hazards and effects. This paper examines the notion of a safety culture in terms of the characteristics of being informed and trusting. These notions are related to more general organisational dimensions describing behaviours and attitudes. Cultures are seen as being defined by their Values, their Beliefs, their Common working Practices and also the ways in which they respond to unusual situations. In a Safety Culture these are all aligned to ensure safe operation even, or especially, when hazardous operations are undertaken. The evolutionary framework of cultures from the Pathological and the Reactive, through the Calculative or Bureaucratic to the Proactive and Generative cultures are described. The Generative culture is equated with the High Reliability Organisations identified in studies of military and civil high risk operations. Next a model is proposed for how to change organisations in order to acquire a safety culture. Finally the barriers to successful intervention are discussed. These include the nature of bureaucratic organisations, the conflicting goals of regulators, failures of management and the fact that change processes are hard.

Introduction - Why Safety Culture?

Safety is an ubiquitous concept. In some industries, such as commercial aviation, safety is so embedded into the organisation that it can be difficult to see just what the general concept of safety means, so I want to start by expanding the notion. Most people see safety as concerned primarily with the personal well-being of stakeholders, by which I mean all those involved, not just the immediate actors and owners. Some also add the integrity of the business and its assets. While these are necessary preconditions, I view safety, and more specifically safety management, in a more active way. I see the creation of a safe environment as allowing dangerous activities to take place successfully, which means without harm or damage. What this means is that safety is more than a passive and well-meaning notion, such as “Thou shall do no harm”. Instead safety is something that has to be actively managed to allow profit or advantage to be gained. The oil and gas industry is one that is naturally dangerous – fire and explosion are natural hazards of the product, mass and power inherent in the means of production. The aviation industry is another; flying is the defiance of gravity and, outside of the Zeppelin, high speeds are also inherent. In both these cases, as well as similar industries, risk is the name of the game. Even an apparently sedentary occupation, such as banking, involves risk and the potential for massive loss. Those organisations that manage their risks best are in place to make the most profit. Those that do not manage so well are either perceived as dangerous or are forced to scale down their operations to achieve acceptable levels of safety.

What has safety culture to do with this? The answer is that there are a number of ways of achieving high levels of safety. These range from having a systematic and highly controlled prescription of all activities in order to exclude the possibility of hazards ever becoming loose, to creating an organisational culture within which everyone is personally involved in ensuring the safety of all concerned, such as DuPont’s *interdependent* culture. The term safety culture can be applied to both, but they clearly represent quite different cultures. What has become clear is that there is a natural and evolutionary progression of cultures, first laid out by Westrum (1991), and that the end-point of this progression is what we call a true Safety Culture. What has also become clear more recently is that, while the road to achieving this ideal state is not an

easy one, the benefits to be gained most certainly outweigh the costs of attaining it. In particular there are advantages to be had from actually reducing the time, and especially paperwork, devoted to safety. The reason for this is that much of what takes place in managing safety in earlier stages is the direct result of a failure of trust and a lack of confidence. These shortcomings lead to over-management and, accordingly, more hard work than is necessary.

This paper first examines the notion of a safety culture and attempts to identify the components in a way that is useful. Then I will discuss how one might go about achieving the goal of being a real safety culture. Next I will discuss briefly some of the barriers that are liable to prevent the full development. Finally I will draw conclusions in which the commercial factor again plays a role.

What is a Safety Culture?

Every organisation has some common, internal, characteristics that we call its culture. These characteristics have often become invisible to those inside, but may be startling to outsiders coming from a different culture. The notion of an organisational culture is notoriously difficult to define (Furnham, 1997; Schein, 1992, 1996), so I take a very general approach and see the organisational culture as, roughly “Who and what we are, what we find important, and how we go about doing things round here”. Rousseau (1988) defined culture more specifically as “the ways of thinking, behaving and believing that members of a social unit have in common”. A safety culture is a special case of such a culture, one in which safety has a special place in the concerns of those who work for the organisation¹.

We can first distinguish culture into its static and its dynamic components. The term *static* refers to what *is*, generally the unchanging values held by the organisation, and the beliefs that permeate its members. The term *dynamic* refers to how the organisation operates, the types of work processes it feels comfortable with. Table 1 shows a set of definitions of the four major components that can be identified as constituting corporate culture (Hudson, 1998). The distinction between common working practices and problem solving methods is not always drawn, but this may be because researchers tend to study companies in either periods of stability or of great change, but not through both. Operating in a stable world highlights the daily working practices, while periods of change are dominated by problem-solving processes. The High Reliability Organisations studied by the Berkeley Group (LaPorte & Consilini, 1991) are characterised by radically different ways of operating under normal and high stress situations.

A safety culture is one in which safety plays a very important role. Because safety is such a complex phenomenon, it is not enough just to add – “And be safe”. The next sections examine the characteristics of a safety culture and look at the types of culture that can be recognised as forming a progression along which organisations develop.

The characteristics of a Safety Culture

What does an organisational culture that gives safety a priority look like? Reason (1997) has identified a number of characteristics that go to make up such a safety culture. These are:

- an *informed culture*-one in which those who manage and operate the system have current knowledge about the human, technical, organisational and environmental factors that determine the safety of the system as a whole,
- a *reporting culture*: a culture in which people are willing to report errors and near misses,
- a *just culture*: a culture of 'no blame' where an atmosphere of trust is present and people are encouraged or even rewarded for providing essential safety-related information- but where there is also a clear line between acceptable and unacceptable behaviour and,

¹ In one sense safety *always* has a place in an organisation's culture, which can then be referred to as *the* safety culture. But it is only past a certain stage of development that an organisation can be said to take safety sufficiently seriously to be labelled as *a* safety culture, a *culture of safety*.

- a *flexible culture* which can take different forms but is characterised as shifting from the conventional hierarchical mode to a flatter professional structure.
- a *learning culture* - the willingness and the competence to draw the right conclusions from its safety information system, and the will to implement major reforms when the need is indicated.

The values associated with a safety culture are fairly straightforward. The beliefs are more complex. Taken together the five characteristics form a culture of *trust* and of *informedness*. Trust is needed, especially in the face of assaults upon the beliefs that people are trying their best, such as accidents and near-miss incidents which all too easily look like failures of individuals to come up to the ideals of the organisation. Informedness means that people know what is really happening, lessening the chance of mistakes caused by inappropriate world views. This helps us to identify what beliefs are associated with a safety culture. Table 2 places safety into the framework set in Table 1. Reason's characteristics are the outcome of corporate behaviours driven by the static and dynamic components of the corporate culture, but mostly by beliefs and behaviours rather than values. Organisations with high values may not live up to their own expectations.

Types of Safety Culture

Safety cultures can be distinguished along a line from *pathological*, caring less about safety than about not being caught, through *calculative*, blindly following all the logically necessary steps, to *generative*, in which safe behaviour is fully integrated into everything the organisation does (Westrum & Adamski, 1999; Westrum, 1991; Weick, 1987). A Culture of Safety can only be considered seriously in the later stages of this evolutionary line. Prior to that, up to and including the calculative stage, the term safety culture is best reserved to describe formal and superficial structures rather than an integral part of the overall culture, pervading how safely the organisation goes about its work. It is obvious that, at the pathological stage, an organisation is not even interested in safety and has to make the first level of acquiring the value system that includes safety as a necessary element. A subsequent stage is one in which safety issues begin to acquire importance, often driven by both internal and external factors as a result of having many incidents. At this first stage of development we can see the values beginning to be acquired, but the beliefs, methods and working practices are still at a primeval stage. At such an early stage, top management believes accidents to be caused by stupidity, inattention and, even, wilfulness on the part of their employees. Many messages may flow from on high, but the majority still reflect the organisation's primary aims, often with '*and be safe*' tacked on at the end. One cannot fail to be 'impressed' by the management of Townsend Thoreson and the messages they were sending to their work force in the run up to the Herald of Free Enterprise disaster (Sheen, 1987).

Table 1: Corporate Culture definitions

<i>Culture Component</i>	<i>Definition</i>
<i>Corporate Values</i>	What the organisation regards as important or even sacrosanct
<i>Corporate Beliefs</i>	What the organisation believes about the world, how the world will react to actions, what the outside world finds important. Beliefs about what works and doesn't
<i>Common Problem-Solving Methods</i>	How the types of problem found in the organisation are tackled, e.g. project groups, consultants, panic
<i>Common Working Practices</i>	The way people go about their work, e.g. small meetings, lots of memos, project management of everything etc.

Table 2: A Safety Culture defined in terms of the organisational components. Note that the methods and working practices are not restricted to safety, but that safety is intimately involved in the way work is done.

<i>Safety Culture Component</i>	<i>Definition</i>
<i>Safety Values</i>	The organisation regards safety as sacrosanct and provides the licence to operate.
<i>Safety Beliefs</i>	The organisation believes that safety makes commercial sense; that individuals are not the sole causes of incidents; that the next accident is waiting to happen.
<i>Common Problem-Solving Methods</i>	Risk assessment, cost-benefit analyses, accident analysis as well as investigation, proactive search for problems in advance of incidents.
<i>Common Working Practices</i>	Safety integral to design and operations practice, safety #1 on meeting agendas up to Board level, chronic unease about safety.

The next stage, one that I feel can not be circumvented, involves the recognition that safety does need to be taken seriously. The term *calculative* is used to stress that safety is calculated; quantitative risk assessment techniques and overt cost-benefit analyses are used to justify safety and to measure the effectiveness of proposed measures. Such techniques are typical problem-solving methods. Often simple calculations suggest that failing to be safe, or at least having incidents, costs money. Furthermore organisations that are seen from outside as being uncaring about safety may have image problems that knock on to the bottom line. Despite this stance, and despite what can become an impressive safety record, safety is still an add-on, certainly when seen from outside.

Table 3: Westrum's original model.

The Reactive and the Proactive stages have been added more recently and articulated in our work in the Oil and Gas industry. Table 5 shows an extended and more practical version that was worked out, in co-operation with Westrum, with the addition of the Reactive and Proactive stages.

Pathological	Bureaucratic	Generative
Information is hidden	Information may be ignored	Information is actively sought
Messengers are "shot"	Messengers are tolerated	Messengers are trained
Responsibilities are shirked	Responsibility is compartmented	Responsibilities are shared
Bridging is discouraged	Bridging is allowed but discouraged	Bridging is rewarded
Failure is covered up	Organisation is just and merciful	Failure causes enquiry
New ideas are crushed	New ideas create problems	New ideas are welcomed

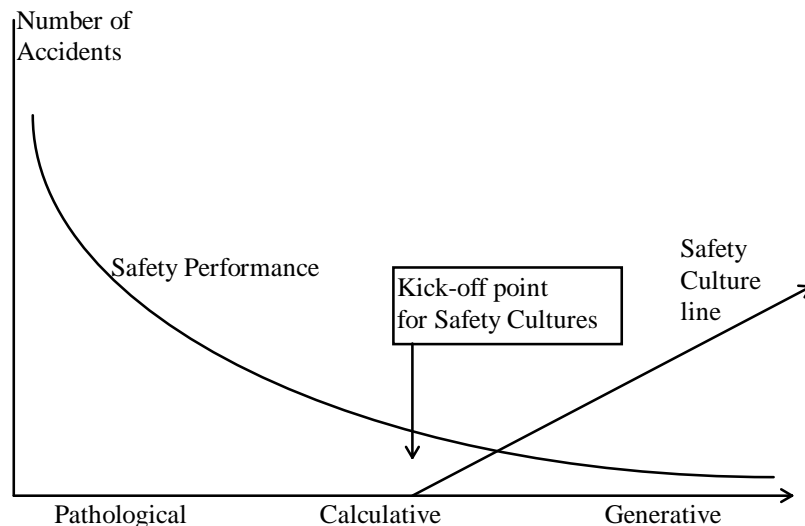


Figure 1: The safety performance will improve as the culture matures, but there can only start to be talk of a Safety Culture once the calculative stage has been passed

The foundation can now be laid, nevertheless, for acquiring *beliefs* that safety is worthwhile in its own right. By constructing deliberate procedures an organisation can force itself into taking safety seriously, or can be forced by a regulatory body, but the values are not yet fully internalised, the methods are still new and individual beliefs generally lag behind corporate intentions. This shows us a significant characteristic of a true safety culture, that the value system associated with safety and safe working has to be fully internalised as beliefs, almost to the point of invisibility, and that the entire suite of approaches the organisation uses are safety-based (Rochlin et al, 1987). What this also stresses is that the notion of a safety culture can only arise in an organisational context within which the necessary technical steps and procedures are already in place and in operation. Yet again, these are necessary but not sufficient preconditions for a safety culture (LaPorte & Consolini, 1991, Laporte, 1996, Turner & Pidgeon, 1997).

Table 4 breaks down general organisational cultures into more detail. The internals may be reflected at any cultural level, so managerial style will vary from pathological through to generative (see below). The Walk/Talk headings are intended to distinguish the more passive from the active components. Filling in these components helps define how a culture appears and how a culture should be. The next section discusses a progression of cultures.

Table 4: The Safety Culture dimensions and internal structure.
These are filled in with different descriptions for each level of the safety culture attained.
For each cell it should be possible to think in terms of the values, beliefs and practices that apply. This is done in Table 5.

TALK		WALK		
Communication	Organisational Attitudes	Safety	Organisational Behaviour	Working Behaviour
Flow of data and information about safety	Workforce attitudes to management	Organisational status of safety Department	Managerial style and behaviour	Priority setting between production and safety
Management informedness about the true state of affairs	Management attitudes about the workforce	Rewards of good safety performance	Level of care for stakeholders	Risk appreciation by those at personal risk
Workforce informedness about the true state of affairs	Collective efficacy – the belief that people can get things done	Procedures and the use of initiative	Dealing with change	On-site behaviour by the workforce and management
		Design – safety as a starting point	Reaction to trouble when it happens	Environment seen as critical

Described in this way we can see how crucial is the notion of belief. The overt knowledge about safety, taken together with a set of values, may still not be enough when difficulties arise, although in easy times behaviour may be exemplary. In the last resort what drives a person, and I would argue an organisation, is less their knowledge than their beliefs. When knowledge clashes with belief the more deep-seated is likely to come out on top as the driver of behaviour, and beliefs, even as articles of unjustified faith, are more deep seated than any rationally acquired knowledge. The latter may be easily disproved or set aside, belief is much harder both to induce and, then, to shift.

The final stages, identified by Westrum in his studies on high reliability organisations and labelled *generative*, involve a much more proactive approach to safety. Whereas the calculative stage represents a reactive approach, using past experience to determine future behaviour, the generative approach may be characterised by a much more internalised model of *good practice* as its driver. This model becomes internalised as a set of beliefs about why and how the organisation operates, about what is the best way to do things. Assumptions about the *need* to be safe are unquestioned; everything else, in contrast, is open for discussion and improvement. A characteristic of this stage is the lack of complacency, even in the face of a dearth of accidents. This has been labelled *chronic unease*, which sums up the pessimistic stance that just because everything has gone well is just an indication that what is about to happen will be a new experience. Fortunately, chronic unease is balanced by optimistic presumption that what *does* happen can be faced and coped with. It does not imply shrinking from challenge, not pessimism elsewhere in the organisation. The generative stage can be equated to the High Reliability Organisations studied by the Berkeley Group (Rochlin et al, 1987; LaPorte & Consolini, 1991; Weick & Roberts, 1993).

One crucial difference between this stage, and prior stages in the evolution of a safety culture, is that the human factor is considered to include both the individual and the organisation. The model of human behaviour has shifted from one in which workers have to be driven, and are not to be trusted, to a more mature understanding of what makes people tick. It is only at this point that it becomes possible to understand that establishment of a safety culture is still not enough, on its own, to counter all human error because such errors may be outside of the control of the immediate perpetrator.

This review suggests that the safety culture concept includes much more than just thinking that safety is important. Work practices and overt priorities not only include safety, but the whole way in which unsafe work is perceived reflects a major shift of point of view. This shift is from regarding individuals as a source of problems for an otherwise perfect organisation to one in which organisations can cause and cure their own problems by using the people who make them up.

Table 5: A more detailed set of descriptions of the different types of safety culture.
HSE is Health, Safety and Environment. This table was defined for the Oil and Gas industry
and has served as a reliable discrimination test.

	PATHOLOGICAL	REACTIVE	CALCULATIVE	PROACTIVE	GENERATIVE
COMMUNICATION	Nobody is informed, no feedback, everybody is passive, no care/ knowledge about safety, don't see(k) or ask the problem, collect what is legally required.	Management demands data on HSE failures, denial until forced to admit, top-down flow of information, bottom-up incidents, lots of statistics nobody understands, safety hot issue after accident.	Environment of command and control by management, lots of HSE graphs, statistics but no follow up, info goes top-down, failures bottom-up, little top-down feedback, toolbox meetings, procedures exist but are only once read. Action is delayed after knowledge.	Management goes out and seek, discuss for themselves they know what to change and how to manage, the feedback loop (bottom-up and top-down) is closing at supervisory level safety topics become part of other meetings, asked for by workforce, they need detail to understand WHY accidents happen.	No threshold between management-workforce, management participates/shares activities (dialogue), HSE is nr 1, all feedback loops are closed, safety is integrated in other meetings; no special safety meetings required, workforce keeps itself up-to-date, they demand information so they can prevent problems.
ORGANISATIONAL ATTITUDES	No believe or trust, environment of punishing, blaming and controlling the workforce.	Failures caused by individuals. No blame but responsibility, workforce needs to be educated and follow the procedures, management overreacts in eyes of workforce.	Workforce is more involved, little effect on procedures, designs, practices workforce does not understand the problem, management is seen as obsessive with HSE, but they don't 'mean' it. (Walk-talk).	Workforce involvement is promoted but ruled/organised by supervisory staff which is obsessed by HSE statistics.	Management is recognised as a partner by workforce, management respects workforce, management has to fix systematic failures, workforce has to identify them.
HSE	No HSE status, HSE issues are ignored, minimal requirements, no rewards on good performance, safety is inherited but not known, reliance on experience.	Meets legal req. collects statistics but no follow up, design is changed after accidents, procedures are rewritten to prevent previous accidents no update or improvements.	HSE well accepted, advisor collects data and creates own statistics, HSE rewards for positive and negative performance, design: quantitative methods, procedures to solve unsolved problems, standard procedures preferred from the shelf, large numbers of procedures but few checks on use/knowledge.	Separate line HSE advisors promoting improvement, but try to reduce the inconvenience to line, for good HSE initiatives there is career enhancement for Sr. staff, HSE is in the early stages of design, procedures are rewritten by workforce, integration with competency, complaints about externally set targets.	HSE department is a small, advising the management on strategy, group, no special rewards, individual pride, procedures are written by workforce, continuous improvement, small numbers of procedures are integrated in training.
ORGANISATIONAL BEHAVIOUR	Denial anything is wrong, avoids HSE discussions, management is hierarchical and stagnant to changes, focus on profits not on workforce, workforce has lots of freedom-> mn don't care.	Man. Holds workforce responsible for failures, overreacting, management. States that it takes safety seriously, but is not always believed by workforce.	Detail focussed/playing with numbers, believe company is doing well in spite of contrary, targets are not challenged, inability to admit solutions may not work the first time.	Management knows the risks, interested in HSE, takes culture into account, safety priority over production which leads to incompatible goals, lots of management walkabouts, communication and assessments about accidents and near-misses and their consequences.	Safety is equal to production, enthusiastic communication between workforce and management and vv, workforce has a lot of freedom-> trust.
WORKING BEHAVIOUR	Workplace is dangerous, messy, no (legal) health requirements, management does not CARE and does not KNOWS.	Basic leg. Requirements implemented, housekeeping is temp. Improved when inspection comes, management KNOWS but not always CARES.	Clean and tidy working environment, housekeeping is very important (prizes), Management CARES but not always KNOWS.	Management CARES and KNOWS, discussion about prioritisation, time and resources are available for sit improvements even before accidents happen.	Management CARES and KNOWS, workforce furnishes its own environment, management passes the experience around to other sites

How can you achieve a Safety Culture?

We have been studying the safety culture of organisations in the oil and gas industry and it is clear that, to progress, one has to undergo a process of cultural change. These changes have to take place incrementally. It appears logical, at least, that it is impossible to go straight from the reactive to the proactive without going through the calculative stage because the proactive culture includes systems typical of the calculative. Similarly it is probably impossible to go from the pathological straight to the calculative stage.

Change Management

What has to be done for an organisation to develop along the line towards the generative or true safety cultures is a managed change process. The next culture defines *where* we want to go to, the change model determines *how* we get there. A model for developmental change has been proposed by Prochaska and DiClemente (1995). This model was originally developed for getting people off drug and other dependencies such as smoking, alcohol and over-eating. It proposes that there are five stages that the authors have identified. These stages are:

- **Precontemplation** – Not yet at a stage of considering the need for change. In safety terms a complacent belief that what can be achieved has been achieved. Coupled with the belief that further improvement is ‘not possible in this business’.
- **Contemplation** – A stage at which the realisation is arisen that further improvement is possible. There is no actual change in behaviour and no steps are taken. Nevertheless the possibility of improvement is entertained.
- **Preparation** – Active steps are taken to prepare for change (in smoking this would be characterised by trying not to buy cigarettes, by not maintaining a stock; in dieting this might involve avoiding certain eating situations, but in both cases without actually smoking or eating less). Characterised by much backsliding.
- **Action** – The stage when the practice built up in the preparation stage is put to work. The beliefs are now that it is important and possible to stop the addictive behaviour. This stage needs to be actively supported while the pull to slide backwards is actively countered (in contrast to the previous stage when backsliding is characteristic).
- **Maintenance** – This stage is vital in maintaining a new, lower baseline of behaviour. This stage needs to be kept up and can often be lost with reversion to the behaviour characteristic of preparation and action.

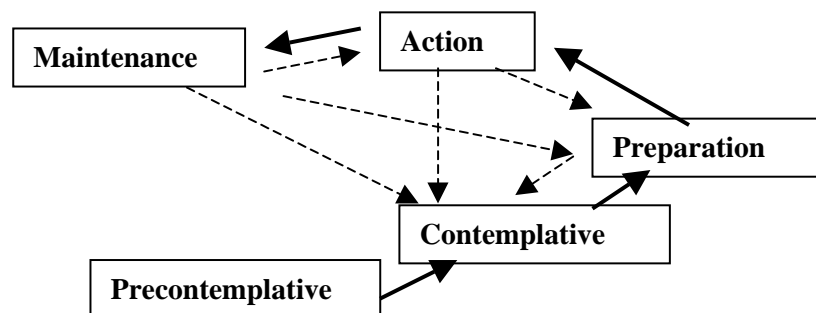


Figure II: Prochaska & DiClemente's change model. The dotted lines denote possible ways to fall back. Note that it is not possible to revert as far as the pre-contemplative mode once one has become aware. The remaining stages are, however, unfortunately quite possible as anyone who has tried to give up smoking knows.

Figure II shows the basic set of transitions from precontemplative through to maintenance, with back-sliding as dotted lines. The step back to precontemplative is not possible (i.e. the values remain intact, but beliefs in the possibility of meeting them may be severely damaged). What is contemplated will be different at each stage of safety culture, so the transition from proactive to generative includes concepts, values and beliefs incomprehensible to those at lower stages. The application of this transition process leads to a spiral when we take safety culture into account.

What is important in this model is the recognition of which stage a patient finds themselves in and the methods available to shift them through the transition from one stage to the next. The stages will require the definition of tools to determine which stage individuals and groups (in organisations) are currently in. The *transitions* that have to be made will require change tools. The term stage is used to refer to one of these treatment situations. A transition takes place between stages.

A Change Model for Organisations

A more articulated model, based upon the simpler Prochaska & DiClemente model, has been developed for managing successful change within organisations rather than individuals. This model, shown below in Table 6, puts together the requirements for change of belief that are so crucial in cultural development. What we have learned is that awareness is not enough, the creation of need and belief in the value of the outcome is equally vital in ensuring a successful process

The model, which has been recently developed in research for Shell International (Hudson et al, 2000), is very similar to any quality system Plan-Do-Check, but the internals of the stages, especially the Awareness and Planning stages, are often missed or treated very summarily. All too often, the active participation of those involved, in the awareness and planning stages, is replaced by a plan of action defined elsewhere. Such plans typically come from senior management, external corporate departments or consultants. What are needed are: (I) the creation of a personal need to change, (II) a belief in the ability to effect such change and (III) the clear understanding that individuals have control over their own process. These are factors that have been repeatedly found in the literature on motivation to influence final outcomes positively. It is just these factors we feel get to the Hearts and Minds of the workforce. When the beliefs and values associated with a new (and hopefully better) state have been assimilated and internalised, then the change has really taken place. This model can apply to safety, but it can also apply to Cost Leadership or any other desirable development in an organisational environment.

Table 6: The articulated Change Model for Organisations. Prochaska and DiClemente's original five stages are elaborated to 14 to cover the details required in real settings.

Pre-contemplation to Contemplation - AWARENESS

- *Awareness* – Simple knowledge of a 'better' alternative than the current state
- *Creation of need* – Active desire to achieve the new state
- *Making the outcome believable* – believing that the state is sensible for those involved
- *Making the outcome achievable*- making the process of achieving the new state credible for those involved
- *Information about successes* - provision of information about others who have succeeded
- *Personal vision* - definition by those involved of what *they* expect the new situation to be

Contemplation to Preparation - PLANNING

- *Plan construction* - creation by those involved of their *own* action plan
- *Measurement points* - definition of indicators of success in process
- *Commitment* - signing-up to the plan of all involved

Preparation to Action - ACTION

- *Do* - start implementing action plan
- *Review* - review progress with concentration upon successful outcomes
- *Correct* - reworking of plan where necessary

Maintenance - MAINTENANCE

- *Review* - management review of process at regular (and defined in advance) intervals
- *Outcome* - checks on internalisation of values and beliefs in outcome state

What are the barriers to success?

If there were no barriers, the development of a safety culture would never form a problem and safety cultures would abound. Why, then, do attempts fail? The reasons are to be found in the beliefs and practices that characterise an organisation and its members. In many cases organisations will naturally limit their development unless active steps are taken. In the worst cases organisations may actually revert. As all organisational cultures past the Pathological hold safety high in their value systems, reversion may appear to the participants to be less significant than it actually is.

Bureaucratic Cultures

One major reason is that the bureaucratic culture associated with the calculative safety culture is a powerful and comfortable one. An organisation that has struggled to become proactive may easily revert, especially in the face of success. Generative organisations have many characteristics that are essentially anti-bureaucratic; the hierarchical structures break down under high-tempo operations (LaPorte & Consilini, 1991). What this demonstrates when it happens is that the beliefs, usually of top management, have never really moved on. The move from proactive to generative is also hard to make because, while the calculative and proactive stages may be fairly easy to identify and therefore acquire, the generative stage is more elusive. In a sense every calculative organisation will be the same, or at least very similar, despite differences in the tasks such organisations face. A generative organisation, in contrast, will be structured in ways specific to the tasks it has to accomplish. Therefore every generative organisation is likely to be subtly different from every other one. This makes it much harder to define where one is going when trying to transit from proactive to generative. It also makes it much easier to succumb to the temptation to prefer a well-defined organisation structure over a process that is much harder to regulate.

Regulators and the Law

The Regulator, possibly surprisingly, forms a barrier to development. This will not be the case in the earlier stages, going from pathological to reactive and on to calculative. The later stages will be harder because they often involve dropping just those facets, such as specialised safety staff and extensive management systems, that regulators require (by law) and that got the organisation there in the first place. Regulators are, with some honourable exceptions, more inclined to the letter than the spirit of the law. This can mean that an experimental improvement, typical of generative and proactive organisations, may well be actively discouraged. The fact that things might well get better is often irrelevant to the legal mind. The simplest remedy for this problem is what is called a goal-setting regime, such as is found in the many offshore oil and gas industries.

The problem faced by an enlightened regulator is that the law allows few distinctions based upon track record in the face of outcomes (Hudson, in prep). What we are looking for is a regulatory regime that is measured against the aspirations of organisations and the degree to which they attempt to attain them. In this sort of regime setting almost impossible standards is laudable, while failing to meet them is not necessarily reprehensible. What counts is the activity and the whole-hearted commitment. In such a regulatory regime meeting low standards might well attract more attention from the regulator than failing to meet high standards. While such enlightened regulatory regimes do not exist, regulators may remain a block to progress by the best.

Management Failure

Changes in top management, or management's priorities, at critical periods, may prove fatal to the successful transition to a higher safety culture. A cultural change is drastic and never takes place overnight. If a champion leaves, there is often no-one to take up the fight and the crucial top-down impetus is lost. But even without a personnel change there are two threats to the successful transition to a higher level of safety culture. One is success, the other failure. In the case of success, effective processes, tools and systems may be dropped, because the problem is perceived to have gone away. In the case of failure, old-fashioned approaches may be retrieved on the grounds that they worked before. But in both of these cases the new, and often fragile, beliefs and practices may not have become sufficiently internalised.

A common problem in organisations that are struggling on the borderline between the calculative and the proactive/generative levels is success. Once significant improvements in outcome performance have been achieved management 'take their eyes off the ball' and downgrade efforts on the grounds that the problems have been solved. But this is behaviour typical of the reactive stance and represents a reversion. Management have to be truly committed to the maintenance of an advanced culture in the face of success, and such commitment is rare.

Change is hard

One underlying reason why cultural change often fails to succeed is that the new situation is unknown to the participants. If this is added to existing beliefs, such as the belief that the current situation is as good as it gets, then there is little real need to change and failure is almost certain. If these failures are at the level of the workforce, then strong management commitment may save the day. If the problems lie with management, then there is little hope because they will enforce the old situation, which feels most comfortable, on the most proactive of workforces. A colleague (G. Old, Pers. Comm.) has likened this to learning a new golf swing by changing the grip and the stance. At first the new position hurts, the old grip position much more comfortable. It takes time before the benefits of a new grip and the altered stance come through, you have to trust the pro, but you have to do the work! One advantage of this metaphor is that managers often play golf and can transfer their experience of learning a new swing to learning to manage an advancing culture. Change agents are like golf professionals, they can help develop a person's game, but they can't play it for them.

Conclusion

The discovery that a safety culture pays is crucial. One way a safety culture pays off, as the levels of trust improve, is in the quality of communication between management, and the rest of the company. As this is always pointed to as a source of problems, having a definitive focus for improving communication can only result in improved performance at all levels. Another way a safety culture pays is in the reduction in time and paperwork devoted to checking whether elementary safety-related actions are carried out. The other main reason why safety makes money lies in the fact that, if one has a guarantee of safety, then one can devote resources more effectively. What costs money is not safety, but bad safety management. Once the management of an organisation realises that safety is financially rewarding and that the costs incurred have to be seen as investments with a positive return (Hudson & Stephens, 2000), the road to a full safety culture should be open.

Given the financial inducements, why don't organisations try and develop the most advanced forms of safety culture? The answer seems to be contained in the type of culture the organisation is at the time. Pathological organisations just don't care. Reactive organisations think that there is nothing better and anyone who claims better performance is probably lying. They do what they feel is as good as can be done. Calculative/Bureaucratic organisations are hard to move because they are comfortable, even if they know that improvement is possible. The more advanced cultures, either Proactive or Generative, are probably easier to attain with small organisations. Large ones will inevitably be heavily bureaucratic unless active steps are taken to counter that tendency.

References

- Furnham, A. (1997) *The Psychology of Behaviour at Work*. Psychology Press, Hove, England
- Hudson, P.T.W. (1998) European Association of Aviation Psychology. Keynote Address, Vienna
- Hudson, P.T.W. (1998) Keynote Address Singapore Aviation Academy
- Hudson P.T.W. & Stephens, D. (2000) Cost and Benefit in HSE: A Model for Calculation of Cost-benefit using Incident Potential. *Proceedings 5th SPE International Conference on Health, Safety and Environment in Oil and Gas Production and Exploration*. CD-ROM, SPE, Richardson, Texas.
- Hudson, P.T.W., Parker, D., Lawton, R., Verschuur, W.L.G., van der Graaf, G.C. & Kalff, J. (2000) The Hearts and Minds Project: Creating Intrinsic Motivation for HSE. . *Proceedings 5th SPE International*

Conference on Health, Safety and Environment in Oil and Gas Production and Exploration. CD-ROM, SPE, Richardson, Texas.

La Porte, T.R. (1996) High Reliability Organizations: Unlikely, Demanding and At Risk. *Journal of Contingencies and Crisis Management*, **4**, 60-71

La Porte, T.R. and Consolini, P.M. (1991) Working in Practice but not in Theory: Theoretical Challenges of High Reliability organizations. *Journal of Public administration Research and Theory*, **1**, 19-47

Prochaska, K. & DiClemente C. (1995) Attitudes to Change. *The American Psychologist*.

Reason, J.T. (1997) *Managing the Risks of Organisational Accidents*. Ashgate, Aldershot.

Rousseau, D. (1988) Quantitative Assessment of Organisational Culture: The Case for Multiple Measures. In L.C. Cooper & I. Robertson (Eds). *International Review of Industrial and Organisational Psychology*. Wiley, Chichester

Sheen. Lord Justice (1987) *The Herald of Free Enterprise*. HMSO, London.

Rochlin, G.I., La Porte, T.R., and Roberts, K.H. (1987) The Self-Designing High-Reliability Organization: Aircraft Carrier Flight Operations at Sea. *Naval War College Review*, **40**, 76-90

Schein, E.H. (1992) *Organizational Culture and Leadership* (2nd Edition). Jossey-Bass, San Francisco.

Schein, E.H. (1996) Culture: The Missing Concept in Organization Studies. *Administrative Science Quarterly*, **41**, 229-240.

Turner, B.A. & Pidgeon, N.F. (1997) *Man-Made Disasters* (2nd Edition). Butterworth Heinemann, Oxford.

Weick, K.E. (1987) Organizational Culture as a Source of High Reliability. *California Management Review*, **29**, 112-127.

Weick, K.E. & Roberts, K.H. (1993) Collective Mind in Organizations: Heedful Interrelating on flight Decks. *Administrative Science Quarterly*, **38**, 357-381.

Westrum, R. (1991) Cultures with Requisite Imagination. In J.Wise, P.Stager & J.Hopkin (Eds.) *Verification and Validation in Complex Man-Machine Systems*. Springer, New York

Westrum, R. & Adamski, A.J. (1999) Organizational Factors Associated with Safety and Mission Success in Aviation Environments. In D.J.Garland, J.A. Wise & V.D.Hopkin (Eds.) *Handbook of Aviation HumanFactors*. Lawrence Erlbaum, Mahwah, NJ.

SHELFS: A Proactive Method for Managing Safety Issues

A. Rizzo, & L. Save

Multimedia Communication Laboratory
University of Siena
Via dei Termini 6
53100 Siena, Italy

Summary. Safety knowledge is an important asset for managing safety critical organisations. In the paper we claim that reactive methods are not the more adequate approach to capture, represent and reuse safety knowledge. The organisational model of accidents and the organisational learning processes ask for a different approach in analysing and documenting safety issues. We present a proactive approach having a holistic view of the productive system, where all the system components and their interactions are analysed. Examples drawn by an experimentation of the method are used to illustrate it.

1. INTRODUCTION

Knowledge is considered the most relevant asset of modern organisations. Most of this knowledge belongs to people and it is embodied in the human practices and interactions among people and artefacts, and it could become organisational knowledge only if properly captured, managed and reused. Modern organisations strive to capture this knowledge since they consider it an important factor for improving the quality of their processes. Yet many safety critical organisations concerning safety issue prefer a reactive approach to learn from experience: the one based on the analysis of reports from accidents, incidents and near misses. Following the direction pointed out by Reason (1991) we claim that reactive methods are not the more adequate approach to capture, represent and reuse safety knowledge. We consider reactive approaches as too slow and inadequate for supporting an efficient experience feedback. Here it is presented a proactive method tailored for introducing human factors in a safety critical company, which is based on a distributed knowledge view of the working processes. This method stresses the positive face of safety and is oriented toward a positive return of experience from the human practices.

Proactive approaches do not just consider events with negative outcomes but also the vital signs of safety, such as the best practices and the solutions identified by managers and operators to overcome organisational and technical problems, and promote the development of such vital signs. Even though this approach was suggested by Reason early in the '90, there are not yet many tools and methods for introducing the approach in large organisations dealing with complex processes with safety critical implication. In addition, there is a lack of methods tailored for organisations that are planning human factors programmes but that do not have a long tradition in human factors. In most of the cases these organisations would like to introduce human factors progressively, having an immediate evidence of the results this introduction is producing. On the contrary, well established and sound methods like HazOp (Kletz, 1993), OARU (Kjellen, & Larrson, 1981) or MORT (Johnson, 1980) require large initial investments, and can be very time-consuming. In addition these methods are not straight oriented to capture best-practices and solutions. Effective organisational learning processes require a return of experience based on an everyday practice involving all the stakeholders involved in a process.

To try to face these issues we present a progressive method oriented toward short-term experience feedback as well as mid and long term actions. The method and related tools aim: 1) at analysing and documenting safety issues for identifying proactive safety actions; 2) at promoting organisational learning as an everyday practice.

In the following we outline a well-know systemic model used to consider the human role in a process and his relationships with the other process components. We elaborated the model on the base of the cultural-historical approach (Cole, 1996) and their recent version known as distributed cognition theory (Norman, 1993) and used the SHEL model as a conceptual framework for developing the method and the tools,

described below. This method have been experimented in a program for introducing human factors principles in the Italian National Railways organisation (FS).

2. THE SHEL MODEL

Edwards (1972) proposed a conceptual model, named SHEL, to describe the behaviour of interactive system with special regard to human factors issues. SHEL is the acronym for Software, Hardware, Environment, and Liveware:

Software represents any component such as the computational code, the policies, norms, rules, procedures, practices and any other formal or informal rules that define the way in which the different components of the system interact among them and with the external environment.

Hardware represents any physical and non-human component of the system as equipment, vehicles, tools, manuals, signs.

Liveware represents any human components in their relational and communicational aspects.

Environment represents the socio-political and economic environment in which the different components of a process interact as shown in Fig. 3.

The SHEL model concentrates on the interfaces among people and all system components including other Liveware resources. SHEL offers a system view where humans cannot be considered isolated from other system components. This view is consistent with a long lasting and empirically well grounded theory of human cognition: the cultural-historical theory, of Vygotsky, Luria and Leontev (for a review see Cole, 1996). Recently, several authors have elaborated along the main ideas of Vygotsky's approach (e. g. Engeström, 1987; Hutchins, 1995; Norman, 1993). The recent elaboration of cultural-historical theory (e.g. Distributed Cognition, Activity Theory) and the SHEL model share the assumption that any productive process is always defined by a specific combination of Hardware, Software and Liveware resources which mediate the execution of human activity. The relationship between the SHEL model and the Vygotsky's unit of analysis of human activity by can be summarized in Figure 1

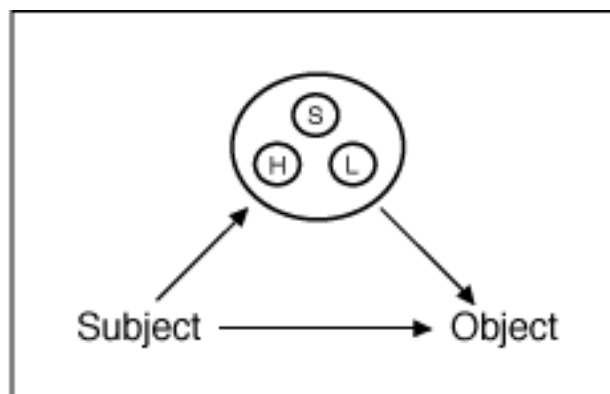


Figure 1: The SHEL model at the light of Vygotsky's unit of analysis of human activity

3. THE SHELFS METHOD AND TOOLS

Using the SHEL model as a possible simplified expression of the Cultural-historical framework we developed a method and relative tools, named SHELFS, for identifying and managing the potential sources of breakdowns in the interaction among human and the other system components. SHELFS was developed within the programme for introducing human factors techniques in the Italian Railways Company (FS). Next sub-session will describe briefly this context of application.

3.1 The context of application

The railways transportation system in Italy is managed by a single organisation named “Ferrovie dello Stato” (FS). Different Departments of FS take care of the railways network, infrastructures, personnel and rolling stock. The “ASA Rete” Department is in charge of the rail tracks management and maintenance and these activities have a direct impact on the safety of the whole rail traffic. Operators involved in rail tracks management usually perform routinely work, in isolated operative contexts. In case of emergency they have to provide quick answers, with few opportunities to verify their decisions with colleagues or with their responsible. Operators of the maintenance section work in teams, usually in hostile environments and under stressing conditions such as the presence of time pressure. Both activities are characterised by the presence of heterogeneous systems interacting with the operators, and by the use of rapidly evolving operative methodologies and technologies. The “ASA Rete” Department identified the need to support the operators involved in these activities, in particular for the aspects of their interactions with the other operators and with the technological and procedural structures they are using. A safety analysis of the organisation evidenced also the need to collect and preserve the safety knowledge of the operators in presence of problems of turnover and downsizing of the company.

As a partial answer to these needs “ASA Rete” launched the “Line Tutor” program. Line tutors are specially selected operators that behave as tutors for their colleagues. They will also analyse the every day operators' activity, under normal and abnormal conditions, with the additional aim of extracting, rationalising and reporting the safety knowledge embedded in their behaviour. Line Tutors have been selected between operators with a well-established experience of the typical operator roles; selection was based on their knowledge and ability for this new position. The SHELFS method was developed for this Line Tutor role, which was supposed to have only a basic knowledge of human factors engineering.

3.2 Method and tools

The method supports the activity of an operator whose role is to identify critical issues and to develop and propose adequate solutions. The method supports also the organised collection, diffusion and re-use of the corporate knowledge existing at the level of single or small group of workers. In particular, corporate knowledge is used during the identification of possible solutions for the critical issues that could originate more serious problems. The operator must have a good knowledge of the working processes and of the working environment he is going to analyse. Approaches concerning “best practice”, as for example the CARMAN approach of Embry and Richardson (1998) or the LINE/LOS checklist of Connelly (1997) shares with SHELFS the aim of documenting safety issues for identifying proactive safety actions. However they are mainly focused on one of the SHEL component, for example the CARMAN approach is a powerful methods to cope with gaps between procedures and practices, and the LINE/LOS checklist is carefully tailored to face Crew communication performance. On the contrary SHELFS try to capture the web of interaction among all the components. Indeed, some of the techniques used in best-practices approaches could be easily integrated into SHELFS, taking for granted that the *distributed cognition* philosophy should drive their application.

The SHELFS method is articulated in three main phases:

- definition of the process;
- identification of the critical issues;
- identification of possible solutions.

In the first phase (definition of the process) the Line Tutor identifies and models the process he is going to analyse. This is done with the direct involvement of the personnel representing each role that is needed to carry out the process. The process is defined with the first tool of the SHELFS method: the Matrix Workflow (see fig 2). The Matrix Workflow allows representing a process according to its basic activities, the personnel involved, the communication flow, the regulations and procedures and the hardware involved.

Process Description Form
















A		Departure from Track 5 of Train BD-813-74 from X to Y									
ACTIVITY		LIVEWARE						SOFTWARE (regulations, directions, procedures.)	HARDWARE (tools, instruments, material, etc.)		
		ROLES INVOLVED									
		MAC	CT	DM	MAN	VER	VEIC				
1	Connecting the engine to the train							Rif. ISM IPCL	Gloves, helmet, lamp		
2	Check connection to the first car							Rif. IPCL			
3	Check brake efficiency (Brake test)							Rif. IEFCA	Hammer, lamp, console manometer		
	Brake Test OK?							B			
4	Delivery of TV40							Rif. IEFCA	TV40		
5	Verbal communication of brake test to the MAC										

Figure 2: The representation of situated process through the workflow matrix

The interactions between humans (Liveware-Liveware) are the element that identifies the different steps in which the process is subdivided: every time the actors change a step is identified, when the actors remain the same also the step remains the same. However, it is always possible to get into the details of a given Liveware-Liveware step by analyzing the interaction between humans and the other components (Liveware-Software and Liveware-Hardware). For example people interactions can be analysed with the conversational model, or the NASA/FAA/LOS checklist (Connelly, 1997); Liveware-Software interactions by checking compliance to procedures; Liveware-Hardware with cognitive walkthrough (Rizzo, et al, 1997).

The output of this phase is a representation of the process under analysis where the focus is on workflow and critical activities of the process itself. Using this representation the Line Tutor can start the second phase (identification of the critical issues) investigating the real breakdowns experienced by the workers while performing the process and the related causes. This is done using a simplified resource analysis method in colloquies and interviews with the workers involved in the process. The resource analysis method is a hierarchical taxonomy that relates the critical issues to the components identified in the SHEL model.

The details of the taxonomy are not very important for the proposed approach, only the 8 main classes of breakdown play an important role.

- H1 Are the tools dependable and effective in playing the role for which they have been introduced in the process?
- H2 The supporting material (e.g. manuals, workbook, signals, etc) supports the activity when needed?
- H3 The physical environment (climate, layout, furniture, etc.) allows a comfortable execution of the activity?
- S1 The knowledge needed to carry out the activity is covered exhaustively by regulations, procedures, instructions, available in the company?
- S2 Practice actually adopted to carry out the activity is consistent with regulations and procedures?
- S3 The specific knowledge needed to carry out the activity is adequate and sufficient?
- L1 The flow of communication is timing and adequate to support the activity?
- L2 The activity distribution, both for the single operator in time and between the operators in time and space, is instrumental to carry out the activity?

Indeed, many of the sub-classes included in the taxonomy are similar to that proposed by other tools as the General Failure Types proposed by Reason, or the Human Error Analytic Taxonomy (Bagnara et al., 1991), or the Project Evaluation Tree put forward by Stephenson (1997). However there are three important differences

with these related works. The first is that human psychophysics conditions (e.g. attention, decision making, reasoning; motivation) are not considered as critical issues since they are strongly influenced by the interactions with all the others system components (Software, Hardware, Liveware) and cannot be faced individually. The second is that using SHELFS the Line Tutor refines the same definition and the analysis of the possible critical issues interactively and iteratively, with the people involved in the process along the three phases of SHELFS. The third is that the three main classes of the taxonomy are not mutually exclusive, on the contrary one critical issue can concern one class as well as all the three main classes. It is important to stress this point since it is at the core of the proposed method. As in the first phase the aim was to map the main classes of resources involved in a process, in this second phase the aim is to assess, according to the experience, how well the resources interact among them.

During this second phase the Line Tutor needs to go only through the 8 main potential critical issues. Three of the critical issues concern the Hardware, three the Software and two the Liveware. The 8 main classes represent different ways of mining the interaction among components. The distinction is not only phenomenological but also grounded in the adopted theoretical approach. Software resources can be prone to wrong interaction since they do not cover all the interaction among components (S1), leaving space for the development of idiosyncratic practices. It is important to note that in complex system, Software resources (e.g. Rules, procedures, computational code, etc.) cannot anticipate all the possible state of components interaction. Notwithstanding this, it is possible to be conscious of this limit and do not pretend that it does not exist. Software can be also not instrumental at good interaction when do not promote the development of working practices consistent with procedures and regulations (S2), or when it do not assure that the relevant knowledge that operator should manage is properly practised in tuition and training (S3). Hardware can embody knowledge that can conflict with Software or Liveware components since degraded, or not anymore adequate to face the evolution of knowledge occurred in the Software and Liveware components, or even since it was not designed at all for the interaction (H1). Hardware can be also prone to faulty interaction when the embodiment of knowledge is carried out with artifacts, like writing or sign devoted to represent other artifacts and modes of action (e.g. manuals, display, signals), which are not tailored to the working condition or since the knowledge representation is not relevant or effective for the interaction (H2). Finally, Hardware can mine the interaction when the physical environment instead to be instrumental to the designed interactions hamper them (H3). The Liveware resource can be fond to mis-interaction when the communication flow, for what concerns both content and form, is fragile and/or not well designed (L1) or when the work distribution among operators and/or in the single operators is not instrumental to the activity (L2).

Notwithstanding the possible lack of attention the organization can have for these sources of potential breakdowns the people involved in the productive system will strive to accommodate them locally, by modifying the relationship between the system components. Sometime this accommodation reveals and creates space for opportunities that should be properly managed by the organization to capture the knowledge they have embedded. The investigation based on SHELFS tends to identify this knowledge and to use it in the identification of solutions for the critical issues (phase 3 of the method).

The aim of the proposed taxonomy is to support the Line Tutor in catching an inadequate distribution of resources for one or more steps of a process. To this aim at least one representative for each of the working positions involved in the process under analysis, is interviewed. This allows the Line Tutor to have a complete idea of all the potential breakdowns associated with that process.

For example, taking into consideration the above reported process “Departure from Track 5 of Train BD-813-74 from X to Y”, in Figure 3 we can observe the summary of the process and the related map of critical issues according to the different roles involved. The critical issues represent a grouping of potential breakdowns, that put together the problems associated with a subset of the whole process and a pool of roles. The critical issue are defined according to the techniques of “one sentence problem statement” (Newman and Lamming, 1995) and in agreement with the operators, which also rate the priority of the critical issues.

INTERVIEW PLAN AND EMERGING CRITICAL ISSUES

	ACTIVITY	MAC 1	MAC 2	MAC 3	MAC 4	CT 1	DM 1	DM 2	VER 1	VER 2	VEI C 1	MA N 1
1	Connecting the engine to the train	90	90									90
2	Check of connection of the first car	90	90									
3	Check of brake efficiency (Brake test)			90	90				90			
Brake Test OK?												
4	Delivery of TV40								90			
5	Verbal communication of brake test to the MAC								90			
6	Verbal communication of brake test to the DM						90		90	90		
7	Delivery of Train Form and M40	90	90				90					
8	Check of Train Form and M40	90	90									
9	Opening of Departure signals					90	90	90				
10	Time schedule check					90						
11	Order of departure					90						
12	RS plug in and train departure											
13	Reaching of maximum speed allowed	90	90									
B1	Communication of excluded car							90	90	90		
B2	Computation of the new of braking mass							90			90	
B3	Delivery of S.T. and new prescription on M40							90			90	

Critical Issues in order of priority:

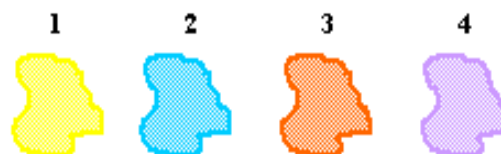


Figure 3: A summary of the critical issues associated with a process.
 The MAC, CT, DM, VER, VEIC and MAN tags represent different roles involved in the process.
 The numeral associated with the tag represent the number of people interviewed.
 The cluster of breakdowns are represented by mean of grey patches.

At the end of this second phase the Line Tutor has a process description with the associated critical issues, and a description of the way in which they are locally managed by re-distributing the Hardware, Software and Human resources.

This constitutes the input for the last phase of the method (identification of possible solutions), where the Line Tutor organises a meeting with the representatives of all the human roles necessary to carry out the process under analysis. The meeting play an important role in the SHELFS approach, it is derived by the participatory meeting proposed by the Scandinavian school (cf. Greenbaun and Kyng, 1991). During this meeting all the critical issues are analysed, discussed and possible solutions are proposed by the same workers involved in the processes under analysis, with the mediating role of the Line Tutor. The meeting (one or more if needed) is organised in four sessions:

- declaration and awareness of critical issues
- critique and analysis of the critical issues
- envisioning solutions
- implementing solutions

In the declaration and awareness session the critical issues collected by the Line Tutor are presented to the participants with the support of the "one sentence problem statement". That is, the Line Tutor summarises in

one sentence a given critical issue reporting in the sentence: the activity; the way in which this activity is hampered; the roles involved; the possible regulations and/or hardware involved.

For example, the first critical issues of the above reported process was “ The Train driver and the Train Manager could not respect/check the maximum speed allowed and reported on the Train Form but not consistent with the maximum speed reported on the M40 form”

The aim of this session is the mutual awareness of the critical issues associated to a given process by all the roles involved. The Output is a list of sentences that express the process critical issues restated and shared by all the participants of the meeting. (see figure 4)



Critical Issue	Activities involved	States of the Critical issues in order	PROBLEM IN ONE SENTENCE
	7-8-9-10 11-13	A	The Train driver and the Train Manager could not respect/check the maximum speed allowed and reported on the Train Form but not consistent with the maximum speed reported on the M40 form
		B	The Train driver and the Train Manager could be constrained to read the travel forms when the Train is already travelling by receiving the documentation not in time or while they were performing other tasks
	3-4-5-6	A	The communication between the Train driver and the Verificator, not adequately supported by the available tools and by the specific training for the role, could produce misunderstanding that delay the departure or do not allow to respect the procedures.
		B	The communication flow among VER, DM, CT, MAC concerning the brake test is not always clear and efficient, with the chance that the Train could departure without that the MAC will become conscious of possible variations on the train characteristics.

Figure 4: Example of “one sentence problem statement” related to the first two critical issues of the “Departure from Track 5 of Train BD-813-74 from X to Y” process

In the Critique and analysis session every critical issues is illustrated by specific events and stories reported by the roles involved. The level of analysis is established according to actions already experimented on the field and according to the interactions among the roles. It is important that the level of analysis of the critical issue will allow the communication between roles even though there can still be substantial differences in the way the problem is perceived. If different levels of analysis are proposed by different roles, the Line Tutor will accept all of them and propose to address the levels one by one. The sentence representing the critical issue is located at the centre of a graph. The details of the criticality, defined according to the SHELFS taxonomy and the roles interested in the critical issue are also represented in the graph, in direct connection with the sentence. The aim of this session is to define the details of the critical issue and the level where it seems manageable. The output is provided by the criticality graphs, which explode a critical issues in relation to the roles involved and the possible factors foreseen by SHELFS.

For example, for the critical issue 1A we had the following S1 and S2 breakdowns:

MAC 1 - The M40 form might disturb me. There are useless prescriptions and other stuff already reported on the Train Form

MAC 2 - The M40 is misleading. If I am tired it can confuse me

MAC 2 -The regulations to which we should refer (REG.243 e PGOS) are warped. In critical situation they can even create big problems.

Which could lead to the best of the case to a violation of the regulation, and in the worse of the case to the overcoming of the maximum speed that a train could safely sustain. according to the class of cars, the percentage of braking mass, etc.

Along this session it was found that the critical issue related to the **Train Form** and to the **M40** form was due to the overlapping of the regulations governing two different type of travel documents: The Train Form itself, recently introduced, and the **Time Pamphlet**, which represent the traditional document supporting the train travel. The Time Pamphlet has always been associated to the M40 as a complementary form where to point out to the Driver possible additional prescription. The same M40 is today associated with the Train Form, which include a more appropriate description of the Train data. This could makes superfluous some of the data reported in the M40 (or viceversa). But why one should remove one of the two prescriptions of speed and which one?

During the Envisioning solution session everyone is free to submit solutions, the only constraint is the request to specify them in relationship to the Software, Hardware and Liveware components. This is a real brainstorming session, so long speech and killing sentences like “this is completely unrealistic” are inhibited by the Line Tutor. The aim of this session is to go away from established position and from defensive and conservative attitude, so to give room to alternative and possible solutions before to prepare the operative proposals.

In the envisioning session it become clear But the real critical issue laid in two different criteria for assessing and reporting the maximum speed, and their relationship with the new philosophy for train traffic management introduced with the new organisation of the FS holding. Here we will not go into the detail of the two regulations, which will need a deep understanding of the work organisation and its history. But we would highlight how the meeting allowed to goes beyond the surface of the problem (apparent redundancy of information). This was fundamental to provide the right rationale for the suggested modifications. Indeed, until the critique and analysis session the two different divisions were blaming each other for the inconvenient: the Train drivers blamed the Train Traffic Manager for providing incorrect prescriptions, instead the Train Traffic Manager blamed the Train Driver to not knowing the rule governing the use of M40. During the meeting both roles devised a shared solution: To eliminate the prescription to report the maximum speed of the train on the M40 if this is higher than the one initially scheduled for that Train. With this solution the Train Driver are not induced in confusion, and the Train Traffic Manager can highlight relevant information in a simpler way. It is important to stress that this apparently simple solution has been accepted only through the shared understanding of the two different criteria for assessing and reporting the maximum speed and their relationship to the new modalities of train traffic management.

In the Implementing solutions session the critical issues are organised by priority, everyone is free to submit his own order and the consensus on priorities in not required. The ranks average decides the order of discussion. The proposals should be feasible in the short/medium term since it is of paramount importance to test them on the field. Moreover, the proposal should specify the possible modalities of implementation and specify the new distribution of knowledge among the Software, Hardware and Liveware components, even if the critical issues is apparently well confined within one component.

The activity of the Line Tutor ends with the implementation of short term actions and their monitoring, and the collection of medium term actions together with the results of the short term actions so to present a deeper analysis for potential improvement of the whole process.

4. CONCLUSION

In experimenting the proposed proactive method we found on average 7 critical issues for each process examined, on average 5 of them where analysed and discussed in the meeting and for 4 of them a shared solution was found. In many cases the critical issues where known to the Line Tutors, but in several other cases the critical issues emerged with the SHELFS method were unknown to the same people involved in the process. For many of them a solution was proposed that could be also extended to other processes that share similar distribution of resources.

REFERENCES

- Bagnara S., Di Martino C., Lisanti B., Mancini G., Rizzo A., 1991, A human error taxonomy based on cognitive engineering and on social and occupational psychology. In *Risk Analysis and safety management in control processes* by G. Apostolakis (Ed) (Academic Press, San Diego), pp. 81-86.
- Cole, M., 1996, *Cultural psychology* (Harvard University Press, Cambridge, MA).
- Connelly, P., 1997, A resource package for CRM developers: Behavioral markers of CRM skill from real world case studies and accidents. Technical Report 97-3. Aerospace Crew Research Project, Department of Psychology, University of Texas, Austin TX 78712.
- Edwards, E., 1972, Man and machine: Systems for safety, *Proceedings of British Airline Pilots Associations Technical Symposium* (British Airline Pilots Associations, London), pp. 21-36.
- Engeström, Y., 1987, *Learning by expanding: An activity-theoretical approach to developmental research* (Orienta-Konsultit, Helsinki).
- Embrey, D., & Richardson, P., 1998, CARMAN: A systematic risk based approach to improving compliance to procedures. Technical Report, Human Reliability Associates, LTD, 1 School House, Higher Line, Dalton, Wigan Lancs WN8 7RP, UK.
- Greenbaun, J., & Kyng, M., 1991, *Design at work: Cooperative design of computer systems* (Lawrence Erlbaum, Hillsdale NJ).
- Hutchins, E., 1995, *Cognition in the Wild* (MIT Press, Cambridge, MA).
- Johnson, W.G., 1980, *MORT Safety assurance systems* (Marcel Dekker, New York).
- Kjellen, U., & Larrson, T.J., 1981, Investigating accidents and reducing risks, *Journal of occupational accidents*, 3, 129-140.
- Kletz, T., 1993, *Lessons from disaster* (Gulf Publishing Company, Houston).
- Newman, W., Lamming, M.G. 1995. *Interactive System Design*. New York: Addison-Wesley.
- Norman, D. A., 1993, *Things that makes us smart* (Addison-Wesley, New York).
- Reason, J., 1991, Too little and too late: A commentary on accident and incident reporting systems. In *Near miss reporting as a safety tool* by van der Schaaf, T.W., Lucas, D.A., Hale, A.R. (Eds.) (Butterworth-Heinemann, Oxford), pp 9-26.
- Rizzo, A., Marchigiani, E., Andreadis, A., 1997, The AVANTI Project: Prototyping and evaluation with a Cognitive Walkthrough based on the Norman's model of action , *Proceedings of Designing Interactive Systems*. (ACM, New York) pp. 305-310.
- Stephenson , J., 1997, *System Safety 2000 : A Practical Guide for Planning, Managing, and Conducting System Safety Programs* (Wiley & Sons, New York).

This page has been deliberately left blank



Page intentionnellement blanche

REPORT DOCUMENTATION PAGE			
1. Recipient's Reference	2. Originator's References RTO-MP-032 AC/323(HFM)TP/12	3. Further Reference ISBN 92-837-1053-3	4. Security Classification of Document UNCLASSIFIED/ UNLIMITED
5. Originator	Research and Technology Organization North Atlantic Treaty Organization BP 25, 7 rue Ancelle, F-92201 Neuilly-sur-Seine Cedex, France		
6. Title	The Human Factor in System Reliability – Is Human Performance Predictable?		
7. Presented at/sponsored by	the Human Factors and Medicine Panel (HFM) Workshop held in Siena, Italy from 1-2 December 1999.		
8. Author(s)/Editor(s) Multiple			9. Date January 2001
10. Author's/Editor's Address Multiple			11. Pages 110
12. Distribution Statement	There are no restrictions on the distribution of this document. Information about the availability of this and other RTO unclassified publications is given on the back cover.		
13. Keywords/Descriptors			
Human factors engineering	Cognition	Adaptation	
Performance	Systems analysis	Design	
Reliability	Aviation safety	Procedures	
Safety	Predictions	Organizations	
Errors	Accident investigations	Models	
Performance evaluation	Humans	Human behavior	
14. Abstract			
<p>Human error is seen as an unacceptably high contributing factor in most military accidents and much research has been carried out over the past 50 years, to attempt to predict the probability of the occurrence of human error. Significant advances have been made within the safety critical domain areas within the nuclear and chemical industries. The aim of the workshop was to review the research carried out across multiple domain areas in order to provide a clear focus for Working Group 30 (Human Reliability in Safety Critical Systems). It was evident from the workshop that key cognitive processes and organisational contexts play an important part in shaping the overall human performance and hence the likelihood of human error. Therefore it was clear that there are new approaches to Human Reliability Assessment that take account of the unique human adaptability attributes that are not present in any other part of the overall system in which the human is an integral part. Working Group 30 will develop these approaches to provide clear guidance to the NATO community in designing and analysing human roles to quantify and qualify the likelihood of error. This will enhance future design processes to produce higher fault tolerant designs, to include mitigating strategies and aim towards a significant reduction in the number of human errors.</p>			

This page has been deliberately left blank



Page intentionnellement blanche



RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25 • 7 RUE ANCELLE

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DIFFUSION DES PUBLICATIONS

RTO NON CLASSIFIEES

L'Organisation pour la recherche et la technologie de l'OTAN (RTO), détient un stock limité de certaines de ses publications récentes, ainsi que de celles de l'ancien AGARD (Groupe consultatif pour la recherche et les réalisations aérospatiales de l'OTAN). Celles-ci pourront éventuellement être obtenues sous forme de copie papier. Pour de plus amples renseignements concernant l'achat de ces ouvrages, adressez-vous par lettre ou par télécopie à l'adresse indiquée ci-dessus. Veuillez ne pas téléphoner.

Des exemplaires supplémentaires peuvent parfois être obtenus auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la RTO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus sur la liste d'envoi de l'un de ces centres.

Les publications de la RTO et de l'AGARD sont en vente auprès des agences de vente indiquées ci-dessous, sous forme de photocopie ou de microfiche. Certains originaux peuvent également être obtenus auprès de CASI.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE

Streitkräfteamt / Abteilung III
Fachinformationszentrum der
Bundeswehr, (FIZBw)
Friedrich-Ebert-Allee 34
D-53113 Bonn

BELGIQUE

Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evère, B-1140 Bruxelles

CANADA

Directeur - Recherche et développement -
Communications et gestion de
l'information - DRDCGI 3
Ministère de la Défense nationale
Ottawa, Ontario K1A 0K2

DANEMARK

Danish Defence Research Establishment
Ryvangs Allé 1, P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE

INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

ETATS-UNIS

NASA Center for AeroSpace
Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320

FRANCE

O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GRECE (Correspondant)

Hellenic Ministry of National
Defence
Defence Industry Research &
Technology General Directorate
Technological R&D Directorate
D.Soutsou 40, GR-11521, Athens

HONGRIE

Department for Scientific
Analysis
Institute of Military Technology
Ministry of Defence
H-1525 Budapest P O Box 26

ISLANDE

Director of Aviation
c/o Flugrad
Reykjavik

ITALIE

Centro di Documentazione
Tecnico-Scientifica della Difesa
Via XX Settembre 123a
00187 Roma

LUXEMBOURG

Voir Belgique

NORVEGE

Norwegian Defence Research
Establishment
Attn: Biblioteket
P.O. Box 25, NO-2007 Kjeller

PAYS-BAS

NDRCC
DGM/DWOO
P.O. Box 20701
2500 ES Den Haag

POLOGNE

Chief of International Cooperation
Division
Research & Development Department
218 Niepodleglosci Av.
00-911 Warsaw

PORTUGAL

Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE

Distribuční a informační středisko R&T
VTÚL a PVO Praha
Mladoboleslavská ul.
197 06 Praha 9-Kbely AFB

ROYAUME-UNI

Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE

Millî Savunma Başkanlığı (MSB)
ARGE Dairesi Başkanlığı (MSB)
06650 Bakanlıklar - Ankara

AGENCES DE VENTE

NASA Center for AeroSpace

Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320
Etats-Unis

The British Library Document

Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume-Uni

Canada Institute for Scientific and

Technical Information (CISTI)
National Research Council
Document Delivery
Montreal Road, Building M-55
Ottawa K1A 0S2, Canada

Les demandes de documents RTO ou AGARD doivent comporter la dénomination "RTO" ou "AGARD" selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications RTO et AGARD figurent dans les journaux suivants:

Scientific and Technical Aerospace Reports (STAR)

STAR peut être consulté en ligne au localisateur de
ressources uniformes (URL) suivant:
<http://www.sti.nasa.gov/Pubs/star/Star.html>
STAR est édité par CASI dans le cadre du programme
NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
Etats-Unis

Government Reports Announcements & Index (GRA&I)

publié par le National Technical Information Service
Springfield
Virginia 2216
Etats-Unis
(accessible également en mode interactif dans la base de
données bibliographiques en ligne du NTIS, et sur CD-ROM)



Imprimé par St-Joseph Ottawa/Hull
(Membre de la Corporation St-Joseph)

45, boul. Sacré-Cœur, Hull (Québec), Canada J8X 1C6



RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25 • 7 RUE ANCELLE

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Telefax 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DISTRIBUTION OF UNCLASSIFIED
RTO PUBLICATIONS

NATO's Research and Technology Organization (RTO) holds limited quantities of some of its recent publications and those of the former AGARD (Advisory Group for Aerospace Research & Development of NATO), and these may be available for purchase in hard copy form. For more information, write or send a telefax to the address given above. **Please do not telephone.**

Further copies are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO publications, or just those relating to one or more specific RTO Panels, they may be willing to include you (or your organisation) in their distribution.

RTO and AGARD publications may be purchased from the Sales Agencies listed below, in photocopy or microfiche form. Original copies of some publications may be available from CASI.

NATIONAL DISTRIBUTION CENTRES

BELGIUM

Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evère, B-1140 Bruxelles

CANADA

Director Research & Development
Communications & Information
Management - DRDCIM 3
Dept of National Defence
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC

Distribuční a informační středisko R&T
VTÚL a PVO Praha
Mladoboleslavská ul.
197 06 Praha 9-Kbely AFB

DENMARK

Danish Defence Research
Establishment
Ryvangs Allé 1, P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE

O.N.E.R.A. (ISP)
29 Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GERMANY

Streitkräfteamt / Abteilung III
Fachinformationszentrum der
Bundeswehr, (FIZBw)
Friedrich-Ebert-Allee 34
D-53113 Bonn

GREECE (Point of Contact)

Hellenic Ministry of National
Defence
Defence Industry Research &
Technology General Directorate
Technological R&D Directorate
D.Soutsou 40, GR-11521, Athens

HUNGARY

Department for Scientific
Analysis
Institute of Military Technology
Ministry of Defence
H-1525 Budapest P O Box 26

ICELAND

Director of Aviation
c/o Flugrad
Reykjavik

ITALY

Centro di Documentazione
Tecnico-Scientifica della Difesa
Via XX Settembre 123a
00187 Roma

LUXEMBOURG

See Belgium

NETHERLANDS

NDRCC
DGM/DWO
P.O. Box 20701
2500 ES Den Haag

NORWAY

Norwegian Defence Research
Establishment
Attn: Biblioteket
P.O. Box 25, NO-2007 Kjeller

POLAND

Chief of International Cooperation
Division
Research & Development
Department
218 Niepodleglosci Av.
00-911 Warsaw

PORTUGAL

Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

SPAIN

INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

TURKEY

Millî Savunma Başkanlığı (MSB)
ARGE Dairesi Başkanlığı (MSB)
06650 Bakanlıklar - Ankara

UNITED KINGDOM

Defence Research Information
Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES

NASA Center for AeroSpace
Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

**NASA Center for AeroSpace
Information (CASI)**

Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320
United States

**The British Library Document
Supply Centre**

Boston Spa, Wetherby
West Yorkshire LS23 7BQ
United Kingdom

**Canada Institute for Scientific and
Technical Information (CISTI)**

National Research Council
Document Delivery
Montreal Road, Building M-55
Ottawa K1A 0S2, Canada

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of RTO and AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)

STAR is available on-line at the following uniform resource locator:

<http://www.sti.nasa.gov/Pubs/star/Star.html>

STAR is published by CASI for the NASA Scientific and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
United States

Government Reports Announcements & Index (GRA&I)

published by the National Technical Information Service
Springfield
Virginia 22161
United States
(also available online in the NTIS Bibliographic Database or on CD-ROM)



Printed by St. Joseph Ottawa/Hull
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada J8X 1C6